Research Article

Empirical Likelihood for Generalized Functional-Coefficient Regression Models with Multiple Smoothing Variables under Right Censoring Data

Hong-Xia Xu,¹ Han-Sheng Zhong,² and Guo-Liang Fan ²

¹Department of Mathematics, Shanghai Maritime University, Shanghai 201306, China

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

Correspondence should be addressed to Guo-Liang Fan; glfan@shmtu.edu.cn

Received 4 December 2019; Revised 12 February 2020; Accepted 25 February 2020; Published 1 April 2020

Academic Editor: Youssef N. Raffoul

Copyright © 2020 Hong-Xia Xu et al. 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.

Empirical likelihood as a nonparametric approach has been demonstrated to have many desirable merits for constructing a confidence region. The purpose of this article is to apply the empirical likelihood method to study the generalized functional-coefficient regression models with multiple smoothing variables when the response is subject to random right censoring. The coefficient functions with multiple smoothing variables can accommodate various nonlinear interaction effects between covariates. The empirical log-likelihood ratio of an unknown parameter is constructed and shown to have a standard chi-squared limiting distribution at the true parameter. Based on this, the confidence region of the unknown parameter can be constructed. Simulation studies are carried out to indicate that the empirical likelihood method performs better than a normal approximation-based approach for constructing the confidence region.

1. Introduction

In studying the relationship between a response and a set of predictor variables or regressors, the mean response variable is often assumed to be a linear regression function of the regressors. Recently, there has been increasing interest and activity in the general area of semiparametric regression modeling in statistics to analyse the high-dimensional data. Semiparametric models are often employed in regression analysis because they well balance between flexibility and fidelity. Among them, scholars have studied the semiparametric functional-coefficient regression model in recent years due to its direct connectivity to the classical linear model. The semiparametric functional-coefficient regression models with multiple smoothing variables have the following form:

$$Y = X^T \beta + Z^T \alpha(U) + \varepsilon, \quad (1)$$

where Y is the response, $X = (X_1, \dots, X_p)^T$, $Z = (Z_1, \dots, Z_d)^T$, and $U = (U_1, \dots, U_d)^T$ are predictors, $\beta =$

$(\beta_1, \dots, \beta_p)^T$ is a vector of p -dimensional unknown parameters, $\alpha(U) = (\alpha_1(U_1), \dots, \alpha_d(U_d))^T$ is a d -dimensional vector of unspecified smooth coefficient functions, and ε is the model error.

Model (1) belongs to the class of semiparametric functional-coefficient regression model and has many useful models as its special cases. Most of previous works are based on the assumption that all coefficient functions share the same smoothing variable. For instance, when all coefficient functions share the same smoothing variable and $\alpha(\cdot) \equiv$ constant, model (1) reduces to the usual linear regression model. When $d = 1$ and $Z \equiv 1$, model (1) reduces to the famous partially linear regression model which was first introduced by Engle et al. [1] to study the influence of weather on electricity demand and further studied by Chen [2], Härdle et al. [3], and Liang and Fan [4] among others. When all coefficient functions share the same smoothing variable, model (1) becomes the varying-coefficient partially linear model which has been widely studied in the literature as well, for instance, the work of Zhang et al. [5], Li et al. [6],

Fan and Huang [7], and Huang and Zhang [8] among others. When the coefficient functions have different smoothing variables, Ip et al. [9] used a generalized likelihood ratio test to test coefficient functions in functional-coefficient regression models; Zhang and Li [10] discussed the functional-coefficient regression models with different smoothing variables in different coefficient functions and defined the integrated estimates of the coefficient functions by marginal integration; Zhang and Li [11] introduced averaged estimation for coefficient functions in functional-coefficient regression models with different smoothing variables; Zhang and Li [12] proposed a procedure for estimating the coefficient functions in the functional-coefficient regression models with different smoothing variables in different coefficient functions; Yang and Lee [13] discussed the semiparametric efficient estimation for generalized functional-coefficient regression models with multiple smoothing variables.

However, the abovementioned related articles are based on the assumption that the data are fully observable. In practice, censored data are widely existing and often encountered. For example, in medical research, termination of studies or subjects leaving studies by various reasons occurs, which is referred to as the random right censoring. The techniques of complete data cannot be directly applied to censored data. Specifically, we consider the following generalized functional-coefficient regression models with multiple smoothing variables:

$$E\{\phi(Y) | X, Z, U\} = X^T \beta + Z^T \alpha(U), \quad (2)$$

where Y is subject to random right censoring and $\phi(\cdot)$ is a unknown function, which is introduced to include various functions according to ones interest, not to deal with random right censoring. For example, it may be the identity function or $I(\cdot \leq t)$, which produces the conditional mean or the conditional probability function given covariates, respectively. When the response was random right censored, Buckley and James [14] gave a method of estimating parameters in the linear regression model; Bravo [15] considered the problem of estimation and inference in semiparametric varying-coefficients partially linear models; for semiparametric varying-coefficient models with different smoothing variables, Yang [16] applied a mean-preserving transformation method to construct estimators for unknown parameters and functions and established their asymptotic normalities.

It is well known that the empirical likelihood (EL) method proposed by Owen [17] has many advantages over the normal approximation (NA) method and the bootstrap methods for constructing confidence intervals. The most appealing features are that the confidence regions constructed by the EL method do not require estimating the covariance of the estimator, and the shape and orientation of confidence regions are only determined by the data. In addition, the confidence region derived by the NA-based method is predetermined to be symmetric which may not be adequate when the underlying distribution is typically asymmetric. More discussions on advantages of the EL

method over the existing methods can be found in the monograph of Owen [18]. In the past two decades, lots of scholars applied the EL approach to investigate various regression models, such as the partially linear model under right censored data [19], partially linear single-index model [20], semiparametric varying-coefficient partially linear regression models [21], longitudinal partially linear model with α -mixing errors [22], high-dimensional partially linear varying-coefficient model with measurement errors [23], linear quantile regression model with the covariates missing at random [24], among others.

In this article, we shall use the EL method to study the generalized functional-coefficient regression models with multiple smoothing variables when the response is subject to random right censoring. The coefficient functions with different smoothing variables can accommodate various nonlinear interaction effects between covariates that are admitted in the model. The empirical log-likelihood ratio for the unknown parameter is constructed based on the smooth backfitting technique, and it is shown to possess a non-parametric version of Wilks' theorem.

The rest of this article is organized as follows. In Section 2, we propose the empirical log-likelihood ratio statistic for the unknown parameter and give the main result. A simulation study is presented in Section 3. Assumption conditions and proofs of the main result are relegated to the Appendix.

2. Methodology

2.1. Introduction for Right Censored Semiparametric Model. Let $W = (X^T, Z^T, U^T)^T$ and C be a set of covariates and a censoring variable, respectively. We assume that a random sample $\{(W_i, Y_i), i = 1, 2, \dots, n\}$ of (W, Y) comes from model (2). In the presence of random right censoring, we observe a random sample (T_i, δ_i) of a pair (T, δ) instead of observing Y_i . The observed Survival Time T is $\min(Y, C)$ and the censoring indicator is $\delta = I(Y \leq C)$, which equals 1 if $Y \leq C$ and 0 otherwise. Since the estimation of the model cannot be obtained directly, the following mean-preserving transformation is considered:

$$Y_G = \frac{\delta \phi(T)}{\{1 - G(T-)\}}, \quad (3)$$

where G is the distribution function of the censoring variable C . In (3), Y_G is zero when the response is censored and $\phi(Y)/\{1 - G(Y-)\}$ when the response is observed. On the other hand, under the assumptions given below, by noting that $E(\delta | W, U) = 1 - G(Y-)$, it follows that

$$E(Y_G | W = w) = E\{\phi(Y) | W = w\}. \quad (4)$$

Although Y_G depends on the choice of ϕ , i.e., $Y_G = Y_G(\phi)$, we omit the dependence in the expression throughout the paper for notational convenience. It also can be seen that (4) holds for any ϕ .

2.2. Empirical Likelihood for Right Censored Data. In this section, we shall apply the EL method to construct

confidence regions of the unknown parameter for generalized functional-coefficient regression models (2) with multiple smoothing variables. For the given β , $\alpha(\cdot)$ can be estimated by using some smoothing techniques with response $Y_{G,i} - X_i^T \beta$. Here, we use the smoothing backfitting technique studied in Lee et al. [25] and Yang [16] and denote the resultant estimator as $\hat{\alpha}_G(\cdot; \beta)$. Let $\hat{\alpha}_{G,j}^{\parallel}(\cdot; \beta)$ satisfy the following system of integral equations:

$$\begin{aligned} \hat{\alpha}_{G,j}^{\parallel}(u_j; \beta) &= \hat{\alpha}_{G,j}^{\parallel}(u_j; \beta) - \sum_{k \neq j} \int \hat{Q}_j(u_j)^{-1} \hat{Q}_{jk}(u_j, u_k) \hat{\alpha}_{G,k} \\ &\cdot (u_k; \beta) du_k, \quad j = 1, \dots, d, \end{aligned} \quad (5)$$

where $\hat{Q}_j(u_j) = n^{-1} \sum_{i=1}^n D_{ij} D_{ij}^T K_{h_j}(u_j, U_{i,j}) Z_{i,j}^2$, $\hat{Q}_{jk}(u_j, u_k) = n^{-1} \sum_{i=1}^n D_{ij} D_{ik}^T K_{h_j}(u_j, U_{i,j}) Z_{i,j} Z_{i,k}$, and $\hat{\alpha}_{G,j}^{\parallel}(u_j; \beta) = n^{-1} \hat{Q}_j(u_j)^{-1} \sum_{i=1}^n D_{ij} K_{h_j}(u_j, U_{i,j}) Z_{i,j} (Y_i^G - X_i^T \beta)$ for $Z_i = (Z_{i,1}, \dots, Z_{i,d})^T$, $U_i = (U_{i,1}, \dots, U_{i,d})^T$, $D_{ij} = \{1, (U_{i,j} - u_j)/h_j\}^T$, and $K_h(u, v) = (K\{(u-v)/h\}) / (\int K\{(w-v)/h\} dw)$ with a base kernel K . Then, $\hat{\alpha}_{G,j}^{\parallel}(\cdot; \beta) = (\hat{\alpha}_{G,1}^{\parallel}(\cdot; \beta)^T, \dots, \hat{\alpha}_{G,d}^{\parallel}(\cdot; \beta)^T)^T$ is called the local linear smooth backfitting estimator with response $Y_i^G - X_i^T \beta$. Note that $\hat{\alpha}_{G,j}^{\parallel}(\cdot; \beta)$ estimates not only the coefficient functions but also their derivatives. Thus, only the first components of $\hat{\alpha}_{G,j}^{\parallel}(\cdot; \beta)$ ($j = 1, \dots, d$) constitute $\hat{\alpha}_G(\cdot; \beta) = (\hat{\alpha}_{G,1}(\cdot; \beta), \dots, \hat{\alpha}_{G,d}(\cdot; \beta))^T$.

An auxiliary random variable is introduced as follows:

$$\eta_i(\beta) = \{X_i - Z_i^T \hat{\alpha}^X(U_i)\} [Y_{G,i} - Z_i^T \hat{\alpha}^{Y_G}(U_i) - \{X_i - Z_i^T \hat{\alpha}^X(U_i)\} \beta]. \quad (6)$$

Similar to Owen [26], an empirical log-likelihood ratio can be defined as

$$\begin{aligned} l_n(\beta) &= -2 \max \left\{ \sum_{i=1}^n \log(n p_i) : \sum_{i=1}^n p_i \eta_i(\beta) = 0, \right. \\ &\quad \left. p_i \geq 0, \sum_{i=1}^n p_i = 1 \right\}. \end{aligned} \quad (7)$$

For any given β , a unique value for $l_n(\beta)$ exists, and we assume that 0 is inside the convex hull of the points $\{\eta_1(\beta), \dots, \eta_n(\beta)\}$. By the Lagrange multiplier method, we can obtain $p_i = \{1 + \lambda^\tau \eta_i(\beta)\}/n$, and $l_n(\beta)$ can be expressed as

$$l_n(\beta) = 2 \sum_{i=1}^n \log\{1 + \lambda^\tau \eta_i(\beta)\}, \quad (8)$$

where λ is the solution of the following equation:

$$\frac{1}{n} \sum_{i=1}^n \frac{\eta_i(\beta)}{1 + \lambda^\tau \eta_i(\beta)} = 0. \quad (9)$$

Theorem 1. Suppose that the conditions (C1)–(C11) in the Appendix hold. If β_0 is the true value of the parameter β , then

$$l_n(\beta_0) \xrightarrow{d} \chi_p^2, \quad \text{as } n \rightarrow \infty, \quad (10)$$

where $\xrightarrow{d}$ stands for convergence in distribution, χ_p^2 is a standard chi-square random variable with p degree of freedom.

It follows from Theorem 1 that an approximate $1 - \alpha$ level confidence region for β_0 can be constructed by $\{\beta: l_n(\beta) \leq c_\alpha\}$, where c_α is chosen to satisfy $P(\chi_p^2 \leq c_\alpha) = 1 - \alpha$.

3. Simulation Study

In this section, we conduct some simulations to assess the performance of the proposed empirical likelihood inference procedure. Throughout this section, the kernel function is taken to be $K(u) = 0.75(1 - u^2)$, $|u| \leq 1$, and the bandwidth h is selected by the cross-validation procedure. We compare our proposed EL method with the normal approximation (NA) method through simulated examples and a real data example.

3.1. Simulated Examples. In this subsection, consider the following model:

$$Y = X_1 \beta_1 + X_2 \beta_2 + Z_1 \alpha_1(U_1) + Z_2 \alpha_2(U_2) + \epsilon, \quad (11)$$

where $\beta_1 = 1.5$, $\beta_2 = 2$, $\alpha_1(U) = 3U^2$, and $\alpha_2(U) = \sin(\pi U)$. The random smoothing variables U_1 and U_2 are independently generated from $U(0, 1)$. The covariates $X = (X_1, X_2)^T$ and $Z = (Z_1, Z_2)^T$ are drawn independently from multivariate normal distribution with zero mean and covariance matrix $(0.5^{|k-l|})$. The model error ϵ is drawn from normal distribution $N(0, 1)$ and the uniform distribution $U(-1, 1)$, respectively.

For random right censoring, we considered $C \sim U(-6, 6) + c_0$, where c_0 is a constant to control the percentage of censoring. In our simulations, we take c_0 as 2.5, 4.0, and 6.6 to yield approximately 30%, 20%, and 10% censoring rate, respectively. We also take $\phi(t) = tI(t \leq t_0)$ for the transformation of the response, where we set t_0 to the largest uncensored observation.

The sample size n is chosen to be 100, 200, and 300, respectively. The coverage probabilities (CPs) and the average lengths (ALs) are calculated based on the EL and NA methods for the normal error and uniform distribution error with 1000 replicates. The simulation results of CPs and ALs with the nominal level $1 - \alpha = 0.95$ are reported in Table 1. Further, we give the confidence regions of (β_1, β_2) with the nominal level $1 - \alpha = 0.95$ in Figures 1 and 2 for the normal error and uniform distribution error, respectively.

Table 1 indicates that the EL method outperforms the NA method, since the EL method gives more accurate coverage probability and shorter average length for almost all settings. We also find that the EL and NA methods perform worse as the censoring rate increases or the sample size decreases. Furthermore, we can see from Figures 1 and 2 that, the EL method performs better than the NA method for different settings. When the censoring rate increases, the confidence region gets larger for two different model error cases. In sum, we strongly recommend our proposed EL

TABLE 1: The coverage probabilities (CPs) and average lengths (ALs) of the confidence intervals based on the EL and NA methods with nominal level $1 - \alpha = 0.95$.

ϵ	n	CR (%)	CP		AL			
			EL	NA	EL (β_1)	EL (β_2)	NA (β_1)	NA (β_2)
$N(0, 1)$	100	10	0.9330	0.9250	0.0249	0.0244	0.0257	0.0255
		20	0.9180	0.9110	0.0270	0.0271	0.0278	0.0283
		30	0.9000	0.8980	0.0281	0.0282	0.0290	0.0291
	200	10	0.9390	0.9330	0.0169	0.0170	0.0172	0.0181
		20	0.9240	0.9210	0.0188	0.0187	0.0210	0.0213
		30	0.9170	0.9150	0.0209	0.0211	0.0228	0.0231
	300	10	0.9440	0.9360	0.0080	0.0090	0.0103	0.0108
		20	0.9380	0.9320	0.0115	0.0123	0.0131	0.0136
		30	0.9260	0.9200	0.0140	0.0133	0.0150	0.0151
$U(-1, 1)$	100	10	0.9290	0.9210	0.0250	0.0248	0.0261	0.0262
		20	0.9070	0.9050	0.0277	0.0276	0.0283	0.0285
		30	0.8980	0.8950	0.0288	0.0286	0.0294	0.0297
	200	10	0.9370	0.9310	0.0171	0.0174	0.0182	0.0180
		20	0.9220	0.9190	0.0190	0.0191	0.0217	0.0216
		30	0.9150	0.9110	0.0214	0.0212	0.0233	0.0239
	300	10	0.9410	0.9350	0.0118	0.0117	0.0123	0.0129
		20	0.9360	0.9300	0.0132	0.0134	0.0140	0.0142
		30	0.9230	0.9140	0.0144	0.0140	0.0152	0.0163

FIGURE 1: Continued.

(c)

FIGURE 1: 95% confidence regions for (β_1, β_2) with CR=10% (a), CR=20% (b), and CR=30% (c) based on $\epsilon \sim N(0, 1)$.

(a)

(b)

FIGURE 2: Continued.

FIGURE 2: 95% confidence regions for (β_1, β_2) with CR=10% (a), CR=20% (b), and CR=30% (c) based on $\epsilon \sim U(-1, 1)$.

FIGURE 3: The confidence regions of (θ_1, θ_2) based on EL and NA methods with nominal level $1 - \alpha = 0.95$ (a) and 0.90 (b), respectively.

method to construct the confidence regions for the unknown parameter in generalized functional-coefficient regression models with multiple smoothing variables under right censoring data.

3.2. Real Data Analysis. In this subsection, we apply our proposed empirical likelihood method to analyse the lung cancer data from the Veteran's Administration lung cancer trial [27]. In this trial, a standard or testing chemotherapy

was randomly assigned to a male suffering from inoperable lung cancer, and 137 males entered into the follow-up at the start of this trial. There are 8 variables for all the patients which were also recorded: Survival in Days (denoted as Y), Status (1 = dead and 0 = censored) (denoted as δ), Treatment Indicator (1 = standard and 2 = test) (denoted as X_1), Cell Type (1 = squamous, 2 = small cell, 3 = adeno, and 4 = large) (denoted as X_2), Age (denoted as U_1), Karnofsky Score (denoted as U_2), Months from Diagnosis (denoted as Z_1), and Prior Therapy (0 = no and 10 = yes) (denoted as Z_2).

Now, we apply model (2) to fit the data and use our proposed EL method and NA method to construct the confidence regions of (β_1, β_2) . The censoring rate of the data is about 6.57%. The profile least squares estimator of β is estimated as $\hat{\beta} = (0.1587, 0.6003)^T$. The confidence regions based on the EL method and NA method of (β_1, β_2) with nominal level $1 - \alpha = 0.95$ and 0.90 are given in Figure 3. From Figure 3, we can see that the confidence regions contain the profile least squares estimate for EL and NA approaches. Furthermore, the confidence regions based on the EL method are smaller than those based on the NA method for different scenarios. Besides, the confidence regions with nominal level $1 - \alpha = 0.95$ are larger than those with nominal level $1 - \alpha = 0.90$. Also, from the confidence regions of the EL method, we can see that the Cell Type is an important factor in the interpretation of Survival in Days. The confidence regions of the Treatment Indicator include zero which indicates that the Treatment Indicator is not a significant factor, that is, there is no difference between the two chemotherapies in the lung cancer treatment.

Appendix

Proof of the Main Results

Denote $\Gamma = E\{[X - \eta(U)^T Z]\{X - \eta(U)^T Z\}^T\}$, $\Delta = E[\sigma_G^2(X, Z, U)\{X - \eta(U)Z\}\{X - \eta(U)Z\}^T]$, $\gamma^k = E(ZZ^T | U = \cdot)^{-1} E(X_k Z | U = \cdot)$, and $\psi^k = \Pi(\gamma^k | \mathcal{H})$, where $\mathcal{H} = \{f: f(u) = (f_1(u_1), \dots, f_d(u_d)), u \in [0, 1]^d, \|f\|^2 < \infty\}$, and $\|f\|^2 = E\{f(U)^T E(ZZ^T | U) f(U)\}$. Here, $\Pi(\mathcal{H})$ denotes the projection onto the space $\mathcal{H}$. In order to prove the main theorems, we first list some regularity conditions which are used in this paper. These conditions are mild and also assumed in [16].

- (C1) Y and C are independent
- (C2) $P(Y \leq C | W, Y) = P(Y \leq C | Y)$
- (C3) $\sup_{t \leq t_0} |\phi(t)| < \infty$ for some t_0 less than the right endpoint of the distribution of the observed time T
- (C4) The probability density function of U is continuous and bounded away from zero on its support $[0, 1]^d$
- (C5) The smallest eigenvalue of $E(ZZ^T | U = u)$ is bounded away from zero on $[0, 1]^d$ and $E(Z_j Z_k | U = u)$ ($j, k = 1, \dots, d$) is continuously partially differentiable in $(0, 1)^d$
- (C6) $\alpha_j(u_j)$ and $\psi_j^k(u_j)$ defined below are twice continuously differentiable in $(0, 1)$
- (C7) $\sup_{u \in [0, 1]^d} E(Z_j^{5+\kappa} | U = u) < \infty$ for $j = 1, \dots, d$ and some $\kappa > 0$ and $E\|X\|^4 < \infty$
- (C8) $E\{ZZ^T \sigma_G^2(X, Z, U) | U = u\}$ is continuous in $[0, 1]^d$, where $\sigma_G^2(X, Z, U) = \text{var}(Y_G | X, Z, U)$
- (C9) The kernel function K is symmetric, Lipschitz continuous in $[-1, 1]$, and has a bounded derivative

(C10) $E\{\exp | X_k - E(X_k | Z, U) | | Z, U\} < \infty$ almost surely for $k = 1, \dots, p$

(C11) The censoring distribution function G is continuous

Lemma A.1. *Under the conditions of Theorem 1, we have $n^{-1/2} \sum_{i=1}^n \eta_i(\beta_0) \xrightarrow{d} N(0, \Delta)$.*

Proof. Observe that

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i(\beta_0) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \{X_i - Z_i^T \hat{\alpha}^X(U_i)\} \\ &\quad \cdot [Y_{G,i} - Z_i^T \hat{\alpha}^{Y_G}(U_i) - \{X_i - Z_i^T \hat{\alpha}^X(U_i)\}^T \beta_0] \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \{X_i - Z_i^T \hat{\alpha}^X(U_i)\} \\ &\quad \cdot \{Y_{G,i} - Z_i^T \hat{\alpha}^{Y_G}(U_i) - X_i \beta + Z_i^T \hat{\alpha}^X(U_i; \beta_0)\} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \{X_i - Z_i^T \hat{\alpha}^X(U_i)\} \\ &\quad \cdot \{Y_{G,i} - X_i^T \beta - Z_i^T \hat{\alpha}_G(U_i; \beta)\}. \end{aligned} \quad (\text{A.1})$$

First, it can be shown that $n^{-1} \sum_{i=1}^n \{X_i - Z_i^T \hat{\alpha}^X(U_i)\}^T \{X_i - Z_i^T \hat{\alpha}^X(U_i)\}^T$ converges in probability to Γ . This comes from the fact that

$$\sup_{u \in [0, 1]^d} |\hat{\alpha}^{X_k}(U) - \psi^k(U)| = O_p\left(h_k^2 + \sqrt{\frac{\log n}{nh_k}}\right), \quad (\text{A.2})$$

and this can be proved from the standard theory of smooth backfitting. Thus,

$$\begin{aligned} \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i(\beta_0) &= n^{-1/2} \sum_{i=1}^n \{X_i - \psi(U_i)^T Z_i\} \epsilon_{G,i} \\ &\quad + n^{-1/2} \sum_{i=1}^n \{\psi(U_i) - \hat{\alpha}^X(U_i)\}^T Z_i \epsilon_{G,i} \\ &\quad + n^{-1/2} \sum_{i=1}^n \{X_i - \hat{\alpha}^X(U_i)^T Z_i\} Z_i^T \{\alpha(U_i) \\ &\quad - \hat{\alpha}_G(U_i; \beta_0)\} := A_{1,n} + A_{2,n} + A_{3,n}. \end{aligned} \quad (\text{A.3})$$

We will prove that $A_{2,n}$ and $A_{3,n}$ are $o_p(1)$. As for $A_{2,n}$, observe that the conditional mean of $A_{2,n}$ given X_i, Z_i, U_i , $i = 1, \dots, n$, is zero, and the conditional variance of k th component of $A_{2,n}$ is given by

$$\begin{aligned} n^{-1} \sum_{i=1}^n \sigma_G^2(X_i, Z_i, U_i) \{\psi^k(U_i) - \hat{\alpha}^{X_k}(U_i)\}^T Z_i Z_i^T \\ \cdot \{\psi^k(U_i) - \hat{\alpha}^{X_k}(U_i)\}. \end{aligned} \quad (\text{A.4})$$

This converges to zero in probability from the fact that

$$\sup_{u \in [0,1]^d} |\hat{\alpha}^{X_k}(U) - \psi^k(U)| = O_p \left(h_k^2 + \sqrt{\frac{\log n}{nh_k}} \right). \quad (\text{A.5})$$

For $A_{3,n} = o_p(1)$, we need to employ an empirical process theory where we need the following properties for $j = 1, \dots, d$:

$$\begin{aligned} \sup_{u_j \in [0,1]} |\hat{\alpha}_{G,j}(U_j; \beta_0) - \alpha_j| &= O_p \left(h_j^2 + \sqrt{\frac{\log n}{nh_j}} \right), \\ \sup_{u_j \in [0,1]} \left| \frac{\partial}{\partial U_j} \hat{\alpha}_{G,j}(U_j; \beta_0) - \frac{\partial}{\partial U_j} \alpha_j(U_j) \right| &= O_p \left(h_j + \sqrt{\frac{\log n}{nh_j^3}} \right). \end{aligned} \quad (\text{A.6})$$

Here, $\hat{\alpha}_{G,j}(u_j; \beta_0)$ and $\alpha_j(\cdot)$ denote the j th component functions of $\hat{\alpha}_G(U_j; \beta_0)$ and $\alpha(\cdot)$, respectively. Since

$$\begin{aligned} A_{3,n} &= n^{-1/2} \sum_{i=1}^n \{X_i - \psi(U_i)^T Z_i\} Z_i^T \{\alpha_0(U_i) - \hat{\alpha}_G(U_i; \beta_0)\} \\ &\quad + n^{-1/2} \sum_{i=1}^n \{\psi(U_i) - \hat{\alpha}^X(U_i)\}^T Z_i Z_i^T \\ &\quad \cdot \{\alpha_0(U_i) - \hat{\alpha}_G(U_i; \beta_0)\}. \end{aligned} \quad (\text{A.7})$$

The second term of the right side of the above equation becomes $o_p(1)$. To show that the first term is $o_p(1)$. By using the empirical theory, we need the exponential tail condition (C10) and also need some moment conditions on $\epsilon_{G,i}$ such as $E|\epsilon_{G,i}|^s$ for $s > (5/2)$. We note that the latter condition is naturally satisfied from the assumptions (C3), (C7), and (C10). Then, it is easy to see that

$$\begin{aligned} A_{1,n} &= o_p(1), \\ A_{2,n} &= o_p(1), \\ A_{3,n} &= o_p(1). \end{aligned} \quad (\text{A.8})$$

Therefore, we have

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i(\beta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{X_i - \psi(U_i) Z_i\} \epsilon_{G,i} + o_p(1). \quad (\text{A.9})$$

It is easy to see that

$$\begin{aligned} \text{var} \left[\frac{1}{\sqrt{n}} \sum_{i=1}^n \{X_i - \psi(U_i) Z_i\} \epsilon_{G,i} \right] &= \frac{1}{n} \sum_{i=1}^n \text{var} \\ &\quad \cdot [\{X_i - \psi(U_i) Z_i\} \epsilon_{G,i}] \\ &= E[\sigma_G^2(X, Z, U)(X - \psi(U)Z)(X - \psi(U)Z)] = \Delta. \end{aligned} \quad (\text{A.10})$$

Hence, by Slutsky's theorem, it follows that $n^{-1/2} \sum_{i=1}^n \eta_i(\beta_0) \xrightarrow{d} N(0, \Delta)$. $\square$

Lemma A.2. Under the conditions of Theorem 1, we have $n^{-1} \sum_{i=1}^n \eta_i(\beta_0) \eta_i^T(\beta_0) \xrightarrow{p} \Sigma$.

Proof. Let $X_i - Z_i^T \hat{\alpha}^X(U_i) = \hat{X}_i$. Observe that

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \eta_i(\beta_0) \eta_i^T(\beta_0) &= \frac{1}{n} \sum_{i=1}^n \{X_i - Z_i^T \hat{\alpha}^X(U_i)\} \{X_i - Z_i^T \hat{\alpha}^X(U_i)\}^T \\ &\quad \cdot [Y_{G,i} - Z_i^T \hat{\alpha}^{Y_G}(U_i) - \{X_i - Z_i^T \hat{\alpha}^X(U_i)\}^T \beta]^2 \\ &= \frac{1}{n} \sum_{i=1}^n \hat{X}_i \hat{X}_i^T Y_{G,i}^2 + \frac{1}{n} \sum_{i=1}^n \hat{X}_i \hat{X}_i^T \{Z_i^T \hat{\alpha}^{Y_G}(U_i)\}^2 \\ &\quad + \frac{1}{n} \sum_{i=1}^n \hat{X}_i \hat{X}_i^T [\{X_i - Z_i^T \hat{\alpha}^X(U_i)\}^T \beta]^2 \\ &\quad - \frac{2}{n} \sum_{i=1}^n \hat{X}_i \hat{X}_i^T Y_{G,i} Z_i^T \hat{\alpha}^{Y_G}(U_i) \\ &\quad - \frac{2}{n} \sum_{i=1}^n \hat{X}_i \hat{X}_i^T Y_{G,i} \{X_i - Z_i^T \hat{\alpha}^X(U_i)\}^T \beta \\ &\quad + \frac{2}{n} \sum_{i=1}^n \hat{X}_i \hat{X}_i^T Z_i^T \hat{\alpha}^{Y_G}(U_i) \{X_i - Z_i^T \hat{\alpha}^X(U_i)\}^T \beta \\ &:= \sum_{i=1}^6 B_{in}. \end{aligned} \quad (\text{A.11})$$

First, we establish $B_{1n} \rightarrow \Sigma$ in probability. For any $p \times 1$ vector α , we have

$$\begin{aligned} E \left(\frac{1}{n} \sum_{i=1}^n \alpha^T \hat{X}_i \hat{X}_i^T \alpha Y_{G,i}^2 \right)^2 &= \frac{1}{n^2} \sum_{i_1=1}^n \sum_{i_2=1}^n (\alpha^T \hat{X}_{i_1})^2 (\alpha^T \hat{X}_{i_2})^2 \\ &\quad \cdot \{E(Y_{G,i_1}^2 Y_{G,i_2}^2) - EY_{G,i_1}^2 EY_{G,i_2}^2\} \\ &= o(1). \end{aligned} \quad (\text{A.12})$$

Thus, we obtain $B_{1n} \rightarrow \Sigma$ in probability. As to $B_{2,n}$, for any $p \times 1$ vector α , we have

$$\begin{aligned} |\alpha^T B_{2,n} \alpha| &= \left| \frac{1}{n} \sum_{i=1}^n \alpha^T \hat{X}_i \hat{X}_i^T \alpha \{Z_i^T \hat{\alpha}^{Y_G}(U_i)\}^2 \right| \\ &\leq \max_{1 \leq i \leq n} \{Z_i^T \hat{\alpha}^{Y_G}(U_i)\}^2 \cdot \frac{1}{n} \sum_{i=1}^n \alpha^T \hat{X}_i \hat{X}_i^T \alpha = o_p(1). \end{aligned} \quad (\text{A.13})$$

Hence, $B_{2,n} = o_p(1)$. As to $B_{3,n}$, for any $p \times 1$ vector α , we have

$$\begin{aligned} |\alpha^T B_{3,n} \alpha| &\leq \left(\max_{1 \leq i \leq n} |\{X_i - Z_i^T \hat{\alpha}^X(U_i)\} \beta| \right)^2 \\ &\quad \cdot \frac{1}{n} \sum_{i=1}^n \alpha^T \hat{X}_i \hat{X}_i^T \alpha = o_p(1), \end{aligned} \quad (\text{A.14})$$

which yields $B_{3,n} = o_p(1)$. Similarly, we can obtain $B_{kn} = o_p(1)$ $k = 4, 5, 6$, which completes the proof of Lemma A.2. $\square$

Lemma A.3. Under the conditions of Theorem 1, we have $\max_{1 \leq i \leq n} \|\eta_i(\beta_0)\| = o_p(n^{1/2})$.

Proof. From Condition (C7) and equation (7) in [16], we have

$$\begin{aligned} \max_{1 \leq i \leq n} \|X_i - Z_i^T \alpha^X(U_i)\| &= o_p(n^{1/4}) \text{ and } \sup_{u_j \in [0,1]} \\ &\cdot \|\hat{\alpha}_{G,j}(u_j; \beta_0) - \alpha_j(u_j)\| \quad (\text{A.15}) \\ &= O_p\left(h_j^2 + \sqrt{\frac{\log n}{nh_j}}\right). \end{aligned}$$

Then, noting that

$$\begin{aligned} \eta_i(\beta_0) &= \{X_i - Z_i \hat{\alpha}^X(U_i)\} [Y_{G,i} - Z_i \hat{\alpha}^{Y_G}(U_i) \\ &\quad - \{X_i - Z_i \hat{\alpha}^X(U_i)\} \beta_0], \end{aligned} \quad (\text{A.16})$$

we can obtain that $\max_{1 \leq i \leq n} \|\eta_i(\beta_0)\| = o_p(n^{1/2})$. $\square$

Lemma A.4. Under the conditions of Theorem 1, we have $\lambda = O_p(n^{-1/2})$.

Proof. Let $\lambda = \rho\theta$, where $\rho \geq 0$, $\theta \in R^p$, and $\|\theta\| = 1$. Set

$$\begin{aligned} \bar{Z} &= \frac{1}{n} \sum_{i=1}^n \eta_i(\beta_0), \\ Z^* &= \max_{1 \leq i \leq n} \|\eta_i(\beta_0)\|, \quad (\text{A.17}) \\ S &= \frac{1}{n} \sum_{i=1}^n \eta_i(\beta_0) \eta_i^T(\beta_0). \end{aligned}$$

From Lemma A.2, we know that S is positive definite when n is large enough. Denote by $\text{mineig}(S)$, the smallest eigenvalues of S . From θ^T times (9), we have

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \frac{\theta^T \eta_i(\beta_0)}{1 + \rho \theta^T \eta_i(\beta_0)} = \frac{1}{n} \sum_{i=1}^n \theta^T \eta_i(\beta_0) \\ &\quad - \rho \frac{1}{n} \sum_{i=1}^n \frac{\{\theta^T \eta_i(\beta_0)\}^2}{1 + \rho \theta^T \eta_i(\beta_0)} \\ &\leq \theta^T \bar{Z} - \frac{\rho}{1 + \rho Z^*} \frac{1}{n} \sum_{i=1}^n \{\theta^T \eta_i(\beta_0)\}^2 = \theta^T \bar{Z} \quad (\text{A.18}) \\ &\quad - \frac{\rho}{1 + \rho Z^*} \frac{1}{n} \sum_{i=1}^n \theta^T \eta_i(\beta_0) \eta_i^T(\beta_0) \theta \\ &\leq \theta^T \bar{Z} - \frac{\rho}{1 + \rho Z^*} \text{mineig}(S), \end{aligned}$$

where we have used $0 < 1 + \lambda^T \eta_i(\beta_0) \leq 1 + \rho Z^*$ which yields from $p_i = (1/n) \sum_{i=1}^n (1/(1 + \lambda^T \eta_i(\beta_0))) \geq 0$. Therefore,

$$\rho \{\text{mineig}(S) - \theta^T \bar{Z} Z^*\} \leq |\theta^T \bar{Z}|. \quad (\text{A.19})$$

By Lemmas A.1 and A.3, we know $Z^* = o_p(n^{1/2})$ and $|\theta^T \bar{Z}| = O_p(n^{-1/2})$. Further, we have

$$\rho \{\text{mineig}(S) + o_p(1)\} \leq |\theta^T \bar{Z}|. \quad (\text{A.20})$$

Lemma A.2 implies $c + o_p(1) \leq \text{mineig}(S) \leq c + o_p(1)$. Hence, $\rho = O_p(n^{-1/2})$ and $\lambda = O_p(n^{1/2})$, which completes the proof of Lemma A.4. $\square$

Proof of Theorem 1. Applying Taylor expansion, from (8) and Lemmas A.1–A.3, we obtain that

$$l_n(\beta_0) = 2 \sum_{i=1}^n \left[\frac{\lambda^T \eta_i(\beta_0) - \{\lambda^T \eta_i(\beta_0)\}^2}{2} \right] + o_p(1). \quad (\text{A.21})$$

From (6), it follows that

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \frac{\eta_i(\beta_0)}{1 + \lambda^T \eta_i(\beta_0)} = \frac{1}{n} \sum_{i=1}^n \eta_i(\beta_0) - \frac{1}{n} \sum_{i=1}^n \eta_i(\beta_0) \eta_i^T(\beta_0) \lambda \\ &\quad + \frac{1}{n} \sum_{i=1}^n \frac{\eta_i(\beta_0) \{\lambda^T \eta_i(\beta_0)\}^2}{1 + \lambda^T \eta_i(\beta_0)}. \end{aligned} \quad (\text{A.22})$$

By Lemmas A.1–A.4, we have

$$\begin{aligned} \left\| \frac{1}{n} \sum_{i=1}^n \frac{\eta_i(\beta_0) \{\lambda^T \eta_i(\beta_0)\}^2}{1 + \lambda^T \eta_i(\beta_0)} \right\| &\leq \frac{1}{n} \sum_{i=1}^n \frac{\|\eta_i(\beta_0)\|^3 \|\lambda\|^2}{|1 + \lambda^T \eta_i(\beta_0)|} \\ &\leq n^{-1} \|\lambda\|^2 \max_{1 \leq i \leq n} \|\eta_i(\beta_0)\| \|\eta_i(\beta_0)\|^2 \\ &= O_p(n^{-1}) o_p(n^{1/2}) O_p(1) = o_p(n^{-1/2}). \end{aligned} \quad (\text{A.23})$$

Then, we find that

$$\begin{aligned} \sum_{i=1}^n \{\lambda^T \eta_i(\beta_0)\}^2 &= \sum_{i=1}^n \lambda^T \eta_i(\beta_0) + o_p(1), \\ \lambda &= \left\{ \sum_{i=1}^n \eta_i(\beta_0) \eta_i^T(\beta_0) \right\}^{-1} \sum_{i=1}^n \eta_i(\beta_0) + o_p(n^{-1/2}). \end{aligned} \quad (\text{A.24})$$

Hence, we have

$$\begin{aligned} l_n(\beta_0) &= \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i(\beta_0) \right\}^T \left\{ \frac{1}{n} \sum_{i=1}^n \eta_i(\beta_0) \eta_i^T(\beta_0) \right\}^{-1} \\ &\quad \cdot \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i(\beta_0) \right\} + o_p(1). \end{aligned} \quad (\text{A.25})$$

Note that $D_1 = \Sigma$, $D_1^{-1}B$, and $B^{1/2}D_1^{-1}B^{1/2}$ have the same eigenvalues. From Lemma A.2, we obtain

$$l_n(\beta_0) = \left\{ B^{-1/2} \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i(\beta_0) \right\}^T B^{1/2} D_1^{-1} B^{1/2} \cdot \left\{ B^{-1/2} \frac{1}{\sqrt{n}} \sum_{i=1}^n \eta_i(\beta_0) \right\} + o_p(1). \quad (\text{A.26})$$

Then, the proof of Theorem 1 is completed. $\square$

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the Project of Humanities and Social Science Foundation of Ministry of Education (20YJC910003), the Key Project of National Social and Scientific Fund Program (18ZDA052), the Project of National Social and Scientific Fund Program (17BGL142), and the Natural Science Foundation of China (91546117).

References

- [1] 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.
- [2] H. Chen, "Convergence rates for parametric components in a partly linear model," *The Annals of Statistics*, vol. 16, no. 1, pp. 136–146, 1988.
- [3] W. Härdle, H. Liang, and J. Gao, *Partially Linear Models*, Physica-Verlag, Heidelberg, Germany, 2000.
- [4] H.-Y. Liang and G.-L. Fan, "Berry-Esseen type bounds of estimators in a semiparametric model with linear process errors," *Journal of Multivariate Analysis*, vol. 100, no. 1, pp. 1–15, 2009.
- [5] W. Zhang, S.-Y. Lee, and X. Song, "Local polynomial fitting in semivarying coefficient model," *Journal of Multivariate Analysis*, vol. 82, no. 1, pp. 166–188, 2002.
- [6] Q. Li, C. J. Huang, D. Li, and T.-T. Fu, "Semiparametric smooth coefficient models," *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 412–422, 2002.
- [7] J. Fan and T. Huang, "Profile likelihood inferences on semiparametric varying-coefficient partially linear models," *Bernoulli*, vol. 11, no. 6, pp. 1031–1057, 2005.
- [8] Z. Huang and R. Zhang, "Empirical likelihood for nonparametric parts in semiparametric varying-coefficient partially linear models," *Statistics & Probability Letters*, vol. 79, no. 16, pp. 1798–1808, 2009.
- [9] W.-C. Ip, H. Wong, and R. Zhang, "Generalized likelihood ratio test for varying-coefficient models with different smoothing variables," *Computational Statistics & Data Analysis*, vol. 51, no. 9, pp. 4543–4561, 2007.
- [10] R. Zhang and G. Li, "Integrated estimation of functional-coefficient regression models with different smoothing variables," *Communications in Statistics - Theory and Methods*, vol. 35, no. 6, pp. 1093–1100, 2006.
- [11] R. Zhang and G. Li, "Averaged estimation of functional-coefficient regression models with different smoothing variables," *Statistics & Probability Letters*, vol. 77, no. 4, pp. 455–461, 2007.
- [12] R. Q. Zhang and G. Y. Li, "Efficient estimation of functional-coefficient regression models with different smoothing variables," *Acta Mathematica Scientia*, vol. 28, no. 4, pp. 989–997, 2008.
- [13] S. J. Yang and Y. K. Lee, "Generalized partially linear varying coefficient models with multiple smoothing variables," *Journal of the Korean Statistical Society*, vol. 43, no. 2, pp. 315–321, 2014.
- [14] J. Buckley and I. James, "Linear regression with censored data," *Biometrika*, vol. 66, no. 3, pp. 429–436, 1979.
- [15] F. Bravo, "Varying coefficients partially linear models with randomly censored data," *Annals of the Institute of Statistical Mathematics*, vol. 66, no. 2, pp. 383–412, 2014.
- [16] S. J. Yang, "Estimation for semiparametric varying coefficient models with different smoothing variables under random right censoring," *Journal of the Korean Statistical Society*, vol. 47, no. 2, pp. 161–171, 2018.
- [17] A. B. Owen, "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, vol. 75, no. 2, pp. 237–249, 1988.
- [18] A. B. Owen, *Empirical Likelihood*, Chapman & Hall/CRC, Boca Raton, FL, USA, 2001.
- [19] Q.-H. Wang and G. Li, "Empirical likelihood semiparametric regression analysis under random censorship," *Journal of Multivariate Analysis*, vol. 83, no. 2, pp. 469–486, 2002.
- [20] L. Zhu and L. Xue, "Empirical likelihood confidence regions in a partially linear single-index model," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 3, pp. 549–570, 2006.
- [21] J. You and Y. Zhou, "Empirical likelihood for semiparametric varying-coefficient partially linear regression models," *Statistics & Probability Letters*, vol. 76, no. 4, pp. 412–422, 2006.
- [22] G. Fan and H. Liang, "Empirical likelihood for longitudinal partially linear model with α -mixing errors," *Journal of Systems Science and Complexity*, vol. 26, no. 2, pp. 232–248, 2013.
- [23] 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.
- [24] Y. Shen, H.-Y. Liang, and G.-L. Fan, "Penalized empirical likelihood for quantile regression with missing covariates and auxiliary information," *Communications in Statistics—Theory and Methods*, vol. 47, no. 8, pp. 2001–2021, 2018.
- [25] Y. K. Lee, E. Mammen, and B. U. Park, "Projection-type estimation for varying coefficient regression models," *Bernoulli*, vol. 18, no. 1, pp. 177–205, 2012.
- [26] A. Owen, "Empirical likelihood ratio confidence regions," *The Annals of Statistics*, vol. 18, no. 1, pp. 90–120, 1990.
- [27] J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, Wiley, Hoboken, NJ, USA, 2nd edition, 2002.