

## Research Article

# Adjusted Empirical Likelihood for Varying Coefficient Partially Linear Models with Censored Data

**Peixin Zhao**

*Department of Mathematics, Hechi University, Yizhou, Guangxi 546300, China*

Correspondence should be addressed to Peixin Zhao; [zpx81@163.com](mailto:zpx81@163.com)

Received 19 August 2012; Revised 19 November 2012; Accepted 24 November 2012

Academic Editor: Xuewen Lu

Copyright © 2013 Peixin Zhao. 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.

By constructing an adjusted auxiliary vector ingeniously, we propose an adjusted empirical likelihood ratio function for the parametric components of varying coefficient partially linear models with censored data. It is shown that its limiting distribution is standard central chi-squared. Then the confidence intervals for the parametric components are constructed. A simulation study and a real data analysis are undertaken to assess the finite sample performance of the proposed method.

## 1. Introduction

Outcome censoring often occurs in many disciplines such as econometrics, biostatistics, and bioinformatics. There have been many recent researches in the area of statistical inference for censored data. Wang and Zheng [1] and Wang and Li [2] considered the estimation problem for partly linear models based on different methods. Yang et al. [3] proposed an empirical likelihood method for a partially linear single-index model with right censored data. More results on the statistical inferences of censored data can be seen in [4–6]. In this paper, we consider the empirical likelihood inferences for a varying coefficient partially linear model with right censored data. Specifically, let  $\{Y_i, i = 1, \dots, n\}$  be the response variable and let  $\{(X_i, Z_i, U_i), i = 1, \dots, n\}$  be associated covariates. We consider the following varying coefficient partially linear model:

$$Y_i = X_i^T \theta(U_i) + Z_i^T \beta + \epsilon_i, \quad (1)$$

where  $\beta$  is a  $q \times 1$  vector of unknown parameters,  $\theta(\cdot)$  is a  $p \times 1$  vector of unknown function, and  $\epsilon_i$  is the model error such as  $E(\epsilon_i | X_i, Z_i, U_i) = 0$ . Due to the curse of dimensionality, as in [7], we assume that  $U_i$  is univariate. In this paper, we assume that the available data  $\{(Y_i^*, \delta_i, X_i, Z_i, U_i), i = 1, \dots, n\}$ , are independent and identically distributed, where  $Y_i^* = \min\{Y_i, C_i\}$ ,  $\delta_i = I(Y_i \leq C_i)$  and  $C_i$  is a censoring

variable. We also assume that  $\{C_i, i = 1, \dots, n\}$  are *i.i.d.* variables with a distribution function  $G(\cdot)$  and  $C_i$  is independent of the prognostic variables  $(X_i, Z_i, U_i)$  and the response  $Y_i$ .

Model (1) has proved to be very useful as it combines the flexibility of nonparametric models and the interpretation of linear models. Recently, a variety of methods have been proposed for estimating the parameters and nonparametric functions in model (1) based on the uncensored data (see [7, 8]). In this paper, we recommend using the empirical likelihood method to make inference for  $\beta$  in the model (1) under right censored data. We define an empirical log-likelihood function based on the synthetic data and show that its asymptotic distribution is a mixture of central chi-squared distribution. Hence, the confidence regions for parametric components can be constructed if the unknown weights are estimated. In addition, by constructing an adjusted auxiliary vector ingeniously, we also propose an adjusted empirical likelihood ratio function for the parametric components and show that its limiting distribution is a standard central chi-squared distribution. Then the confidence intervals for parametric components are constructed and the estimation for the unknown weights is avoided. To demonstrate the performance of the proposed method, based on some numerical experiments such as the simulation study and the real data analysis, we conducted a comparison with the normal approximation method. Compared with the normal approximation method, the confidence intervals based on the empirical likelihood method perform fairly well.

The confidence intervals obtained by the empirical likelihood method possess several attractive features compared to the conventional Wald-type confidence intervals, such as circumvention of asymptotic variance estimation to compute the standard error, flexible shapes of the confidence intervals determined by data and range-preserving property (see [9]). Wang and Li [2] considered the adjusted empirical likelihood inferences for a class of partially linear models with right censored data. Yang et al. [3] considered the adjusted empirical likelihood inferences for a partially linear single-index model with right censored data. This paper also contributes to the rapidly growing literature on the adjustment-based empirical likelihood method and provides additional positive results for the Wilks phenomena in the parametric components of varying coefficient partially linear models with censored data, which extends the application literature of the adjusted empirical likelihood method. The more works on empirical likelihood methods can be found in [10, 11].

## 2. Methodology and Results

As in [3], let  $Y_{iG}^* = Y_i^* \delta_i / (1 - G(Y_i^*))$ ; then we have  $E(Y_{iG}^* | X_i, Z_i, U_i) = E(Y_i | X_i, Z_i, U_i)$ . This implies that

$$Y_{iG}^* = X_i^T \theta(U_i) + Z_i^T \beta + \epsilon_i^*, \quad (2)$$

where  $E(\epsilon_i^* | X_i, Z_i, U_i) = 0$ . When  $G(\cdot)$  is known as well, model (2) is a standard semiparametric varying coefficient partially linear model. For given  $\beta$ , using the same arguments as Fan and Huang [7], we can get the weighted local least-squares estimator of  $\theta(u)$  by minimizing

$$\sum_{i=1}^n \left\{ Y_{iG}^* - Z_i^T \beta - \sum_{k=1}^p [a_k + b_k(U_i - u)] X_{ik} \right\}^2 K_h(U_i - u), \quad (3)$$

where  $K_h(\cdot) = h^{-1}K(\cdot/h)$ ,  $K(\cdot)$  is a kernel function,  $h$  is a bandwidth, and  $X_{ik}$  denotes the  $k$ th component of  $X_i$ .

Let  $Z = (Z_1, \dots, Z_n)^T$ , let  $Y_G^* = (Y_{1G}^*, \dots, Y_{nG}^*)^T$ , let  $I_p$  be  $p \times p$  identity matrix, let  $0_p$  be  $p \times p$  zero matrix, let  $\Omega_u = \text{diag}(K_h(u - U_1), \dots, K_h(u - U_n))$  be  $n \times n$  diagonal matrix. Then, the solution to minimize (3) can be given by

$$\bar{\theta}(u) = (I_p, 0_p) (D_u^T \Omega_u D_u)^{-1} D_u^T \Omega_u (Y_G^* - Z\beta), \quad (4)$$

where  $D_u = \begin{pmatrix} X_1 & \dots & X_n \\ h^{-1}(U_1 - u)X_1 & \dots & h^{-1}(U_n - u)X_n \end{pmatrix}^T$ .

Let  $(I_p, 0_p)(D_u^T \Omega_u D_u)^{-1} D_u^T \Omega_u \equiv (S_1(u), \dots, S_n(u))$ , let  $\bar{\mu}(u) = \sum_{k=1}^n S_k(u)Z_k^T$ , and let  $\bar{g}(u) = \sum_{k=1}^n S_k(u)Y_{kG}^*$ . Then we have  $\bar{\theta}(u) = \bar{g}(u) - \bar{\mu}(u)\beta$ . Substituting this into (2) and by a simple calculation we have

$$\check{Y}_{iG}^* = \check{Z}_i^T \beta + \epsilon_i^*, \quad (5)$$

where  $\check{Z}_i = Z_i - \bar{\mu}(U_i)^T X_i$ ,  $\check{Y}_{iG}^* = Y_{iG}^* - \bar{g}(U_i)^T X_i$ . To construct the empirical likelihood ratio function for  $\beta$ , we introduce the auxiliary random vector

$$\check{\eta}_i(\beta) = \check{Z}_i (\check{Y}_{iG}^* - \check{Z}_i^T \beta). \quad (6)$$

By (5) we have that  $\beta$  is true if and only if  $E\{\check{\eta}_i(\beta)\} = o(1)$ . Using such information, we can define a profile empirical log-likelihood ratio function for  $\beta$ . However,  $G(\cdot)$  is usually unknown, and  $\check{\eta}_i(\beta)$  cannot be used directly to make inference for  $\beta$ . To solve this problem, we replace  $G(\cdot)$  in  $\check{\eta}_i(\beta)$  by its estimator. In this paper, we employ the Kaplan-Meier estimator

$$G_n(y) = 1 - \prod_{i=1}^n \left( \frac{N^+(Y_i^*)}{1 + N^+(Y_i^*)} \right)^{I(Y_i^* \leq y, \delta_i = 0)}, \quad (7)$$

where  $N^+(y) = \sum_{j=1}^n I(Y_j^* > y)$ . Hence, an estimator of  $\check{\eta}_i(\beta)$  can be defined as

$$\hat{\eta}_i(\beta) = \check{Z}_i (\check{Y}_{iG_n}^* - \check{Z}_i^T \beta), \quad (8)$$

where  $\check{Y}_{iG_n}^* = Y_{iG_n}^* - \hat{g}(U_i)^T X_i$  and  $\hat{g}(u) = \sum_{k=1}^n S_k(u)Y_{kG_n}^*$ . Then, we can define the profile empirical log-likelihood ratio function for  $\beta$  as

$$\hat{R}(\beta) = -2 \max \left\{ \sum_{i=1}^n \log(np_i) \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \hat{\eta}_i(\beta) = 0 \right\}, \quad (9)$$

where  $p_i = p_i(\beta)$ ,  $i = 1, \dots, n$ . A unique value for  $\hat{R}(\beta)$  exists, provided that 0 is inside the convex hull of the points  $(\hat{\eta}_1(\beta), \dots, \hat{\eta}_n(\beta))$ . Based on the method of Lagrange multipliers to find the optimal  $p_i$ , then the empirical log-likelihood ratio function can be represented as

$$\hat{R}(\beta) = 2 \sum_{i=1}^n \log(1 + \lambda^T \hat{\eta}_i(\beta)), \quad (10)$$

where  $\lambda = \lambda(\beta)$  is a  $q \times 1$  vector, which satisfies

$$\sum_{i=1}^n \frac{\hat{\eta}_i(\beta)}{(1 + \lambda^T \hat{\eta}_i(\beta))} = 0. \quad (11)$$

Let

$$\Psi(u) = E(XX^T | U = u), \quad \Phi(u) = E(XZ^T | U = u),$$

$$H(s) = \frac{E\{[Z - \mu(U)^T X] Y_G^* I(s < Y^*)\}}{(1 - G(s))(1 - F(s-))},$$

$$\mu(u) = \Psi(u)^{-1} \Phi(u), \quad \Lambda^G(u) = \int_{-\infty}^u \frac{1}{1 - G(s-)} dG(s),$$

$$\Sigma_0(\beta) = E\left\{ [Z - \mu(U)^T X] [Z - \mu(U)^T X]^T \times [Y_G^* - Z^T \beta - X^T \theta(U)]^2 \right\},$$

$$\Sigma_1 = \int_{-\infty}^{\infty} H(s) H(s)^T (1 - F(s-)) (1 - \Delta \Lambda^G(s)) dG(s). \quad (12)$$

The following theorem gives the asymptotic distribution of  $\hat{R}(\beta)$ .

**Theorem 1.** Suppose that conditions (C1)–(C5) in the Appendix hold. If  $\beta_0$  is the true parameter, then

$$\widehat{R}(\beta_0) \xrightarrow{\mathcal{L}} w_1 \chi_{1,1}^2 + w_2 \chi_{1,2}^2 + \cdots + w_q \chi_{1,q}^2, \quad (13)$$

where  $w_1, \dots, w_q$  are the eigenvalues of  $\Sigma_0^{-1}(\beta_0)\Sigma(\beta_0)$ ,  $\Sigma(\beta_0) = \Sigma_0(\beta_0) - \Sigma_1$ , and  $\chi_{1,1}^2, \dots, \chi_{1,q}^2$  are independent standard chi-square random variables with 1 degree of freedom.

From Theorem 1, we can see that the asymptotic distribution of  $\widehat{R}(\beta)$  is a mixture of central chi-square distributions. Hence, the confidence regions for the parametric components can be constructed if the unknown weights are estimated. Next, we give an adjusted empirical log-likelihood ratio function that has an asymptotic standard chi-square distribution. Then the confidence regions for the parametric components are constructed, and the estimation for the unknown weights is avoided. Let  $F_n(\cdot)$  be the Kaplan-Meier estimator for the distribution function of  $Y$ , and

$$\widehat{\Sigma}_0(\beta) = \frac{1}{n} \sum_{i=1}^n \widehat{\eta}_i(\beta) \widehat{\eta}_i(\beta)^T,$$

$$\Lambda_n^{G_n}(u) = \int_{-\infty}^u \frac{1}{1 - G_n(s-)} dG_n(s),$$

$$H_n(s) = \frac{n^{-1} \sum_{i=1}^n \left[ Z_i - \bar{\mu}(U_i)^T X_i \right] Y_{iG_n}^* I(s < Y_i^*)}{(1 - G_n(s))(1 - F_n(s-))},$$

$$\widehat{\Sigma}_1 = \frac{1}{n} \sum_{i=1}^n (1 - \delta_i) H_n(Y_i^*) H_n(Y_i^*)^T (1 - \Delta \Lambda_n^{G_n}(Y_i^* -))$$

$$\widehat{\Sigma}(\beta) = \widehat{\Sigma}_0(\beta) - \widehat{\Sigma}_1, \quad S_n(\beta) = \sum_{i=1}^n \widehat{\eta}_i(\beta) \sum_{i=1}^n \widehat{\eta}_i(\beta)^T. \quad (14)$$

Then, as in [3], an adjusted empirical log-likelihood ratio function can be defined by

$$\widehat{R}_{\text{ad}}(\beta) = \gamma_n(\beta) \widehat{R}(\beta), \quad (15)$$

where  $\gamma_n(\beta) = \text{tr}\{\widehat{\Sigma}^{-1}(\beta)S_n(\beta)\} / \text{tr}\{\widehat{\Sigma}_0^{-1}(\beta)S_n(\beta)\}$ . The following theorem shows that the asymptotic distribution of  $\widehat{R}_{\text{ad}}(\beta)$  can be approximated by standard chi-square distribution.

**Theorem 2.** Suppose that conditions (C1)–(C5) in the Appendix hold. If  $\beta_0$  is the true parameter, then

$$\widehat{R}_{\text{ad}}(\beta_0) \xrightarrow{\mathcal{L}} \chi_q^2. \quad (16)$$

Let  $\chi_q^2(1 - \alpha)$  be the  $1 - \alpha$  quantile of  $\chi_q^2$  ( $0 < \alpha < 1$ ). Theorem 2 implies that an approximate  $1 - \alpha$  confidence region for  $\beta$  can be defined by

$$C_\alpha(\beta) = \{\beta \mid \widehat{R}_{\text{ad}}(\beta) \leq \chi_q^2(1 - \alpha)\}. \quad (17)$$

### 3. Numerical Results

In this section, we conduct several simulation experiments to illustrate the finite sample performances of the proposed method and consider a real data set analysis for further illustration.

**3.1. Simulation Studies.** To evaluate the performance of the proposed adjustment-based empirical likelihood (AEL) method, we present some simulation results. In the simulation, we simulated data from the following model:

$$Y = (2U^2 - 1)X + Z + \epsilon, \quad (18)$$

where  $\beta = 1$ ,  $\theta(u) = 2u^2 - 1$ . We generated  $n = 50, 100$  and  $150$  subjects, respectively, and the covariates  $X \sim N(0, 1)$ ,  $Z \sim N(1, 1.5)$ , and  $U \sim U(-1, 1)$ .  $Y$  is generated according to the model with  $\epsilon \sim N(0, 0.5)$ . The censoring variable  $C \sim N(\mu, 1)$ , where  $\mu = 3.5, 2.1$  and  $1.3$ , respectively, such that the corresponding censoring rates (CR) are about  $0.1, 0.3$ , and  $0.45$ . We use the Epanechnikov kernel function  $K(u) = 0.75(1 - u^2)_+$ , and the cross-validation bandwidth  $h_{\text{CV}}$  is obtained by minimizing

$$\text{CV}(h) = \sum_{i=1}^n (Y_{iG_n}^* - X_i^T \widehat{\theta}_{[i]}(U_i) - Z_i^T \widehat{\beta}_{[i]})^2, \quad (19)$$

where  $\widehat{\beta}_{[i]}$  and  $\widehat{\theta}_{[i]}(\cdot)$  are estimators of  $\beta$  and  $\theta(\cdot)$  which are computed with all of the measurements, but the  $i$ th subject is deleted.  $\widehat{\beta}_{[i]}$  can be obtained by minimizing (10),  $\widehat{\theta}_{[i]}(\cdot)$  can be obtained by replacing  $G(\cdot)$ , and  $\beta$  in (4) by  $G_n(\cdot)$  and  $\widehat{\beta}_{[i]}$ , respectively.

For comparison, we consider two approaches for constructing the confidence intervals: the AEL method proposed in this paper and normality approximation (NA) method proposed by [12]. To construct confidence intervals based on the NA method, we need to estimate the asymptotic variance. However, note that the asymptotic variance of the estimator has a complicated structure; then we estimate the asymptotic variance by the bootstrap method. The average lengths and coverage probabilities of the confidence intervals are shown in Table 1. Here the nominal level is taken as  $1 - \alpha = 95\%$ , and the simulation is computed with 1000 simulation runs. From Table 1, we can see the following observations.

- (i) For any given level of CR and the size of sample, although the coverage probabilities of the confidence intervals obtained by the AEL method and the NA method are similar, the lengths of the confidence intervals obtained by the AEL method are shorter than those obtained by the NA method. This implies that the confidence intervals obtained by the AEL method outperform those obtained by the NA method.
- (ii) For the given size of sample  $n$ , the performances of the confidence intervals, obtained by the AEL method under the moderate censoring rate, are all close to the nominal level 95%. This implies that the adjustment

scheme is workable for the moderate censoring rate. However, the simulation results also suggest that larger samples would be needed when the censoring rate is relatively high such as CR = 0.45.

**3.2. Application to CGD Data.** We illustrate the applicability of our proposed method using the chronic granulomatous disease (CGD) data set from the International CGD Cooperative Study Group. This data set was designed to have a single interim analysis when the follow-up data as of July 15, 1989 were complete. The monitoring committee for the trial terminated the trial at a meeting on September 22, 1989. The treatment given each patient was unblinded at the first scheduled visit for the patient following the decision of the monitoring committee. More details for this data description and analysis can be seen in [13].

The variables contained here are  $X$ : treatment code, 1 = rIFN, 2 = placebo;  $Z_1$ : pattern of inheritance, 1 = X-linked, 2 = autosomal recessive;  $Z_2$ : age, in years;  $Z_3$ : height, in cm;  $Z_4$ : weight, in kg;  $Z_5$ : using corticosteroids at time of study entry, 1 = yes, 2 = no;  $Z_6$ : using prophylactic antibiotics at time of study entry, 1 = yes, 2 = no;  $Z_7$ : 1 = male, 2 = female;  $Z_8$ : hospital category, 1 = US-NIH, 2 = US-other, 3 = Europe, Amsterdam, 4 = Europe, other;  $T$ : elapsed time (in days) from randomization to diagnosis of a serious infection or if a censored observation: elapsed time from randomization to censoring date;  $\delta$ : censoring indicator, 1 = Noncensored observation, 2 = censored observation.  $S$ : sequence number. For each patient, the infection records are in sequence number order.

Based on the generalized likelihood ratio testing, Jiang and Qian [14] show that this data set can be fitted by the following model:

$$Y = X\theta(U) + \sum_{k=0}^8 Z_k \beta_k + \varepsilon, \quad (20)$$

where  $Y = \ln(T)$ ,  $U = S/20$ , and  $Z_0 = 1$  as the intercept term. The estimators and 95% confidence intervals are shown in Table 2. From Table 2, we see that the covariates  $Z_1$ ,  $Z_5$ , and  $Z_7$  have somewhat stronger associations with the response. In addition, we also can see that the confidence intervals obtained by the AEL method are workable.

## Appendix

### A. Proof of Theorems

For convenience and simplicity, let  $C$  denote a positive constant which may be different value at each appearance throughout this paper. Before we prove our main theorems, we list some regularity conditions which are used in this paper.

- (C1) The bandwidth  $h = Cn^{-1/5}$ , for some constant  $C > 0$ . The kernel  $K(\cdot)$  is a symmetric probability density function, and  $\int u^4 K(u) du < \infty$ .
- (C2)  $\theta(u)$ ,  $\sigma^2(u)$ ,  $\Phi(u)$ , and  $\Psi(u)$  are twice continuously differentiable on  $(0, 1)$ , where  $\sigma^2(u) = E(\varepsilon^2 | U = u)$ .

(C3) The density function of  $U$ , says  $f(u)$ , is bounded away from 0 and infinity on  $[0, 1]$  and is continuously differentiable on  $(0, 1)$ .

(C4) For  $s \leq \tau_Q = \inf\{y : Q(y) = 1\}$ ,  $G(s)$  and  $F(s)$  do not have common jumps, where  $Q(y) = P(Y^* \leq y)$ . Moreover, we assume that

$$\int_0^{\tau_Q} \|H(Y)\|^2 (1 - \Lambda^G(s)) dG(s) < \infty, \quad (A.1)$$

$$E \left\{ \frac{\|Z - \mu(U)\| |Y|}{[(1 - G(Y))(1 - F(Y))]^{1/2}} \right\} < \infty.$$

(C5) For given  $u$ ,  $\Psi(u)$  is a positive definite matrix, and  $E\{[Z - \mu(U)^T X][Z - \mu(U)^T X]^T\}$  is nonsingular.

The proofs of theorems rely on the following lemmas.

**Lemma 3.** Let  $(X_1, Y_1), \dots, (X_n, Y_n)$  be i.i.d. random vectors, where  $Y_i$  is scalar random variables. Assume that  $\sup_x \int |y|^s f(x, y) dy < \infty$  and  $E|Y_1|^s < \infty$ , where  $f(\cdot, \cdot)$  denotes the joint density of  $(X, Y)$ . Let  $K(\cdot)$  be a bounded positive function with a bounded support, satisfying a Lipschitz condition. Then

$$\sup_x \left| \frac{1}{n} \sum_{i=1}^n \{K_h(X_i - x) Y_i - E[K_h(X_i - x) Y_i]\} \right| \quad (A.2)$$

$$= O_p \left( \left\{ \frac{\log(1/h)}{nh} \right\}^{1/2} \right),$$

provided that  $n^{2\delta-1}h \rightarrow \infty$ , for some  $\delta < 1 - s^{-1}$ .

*Proof.* This follows immediately from the result that was obtained by Mack and Silverman [15].  $\square$

**Lemma 4.** Suppose that conditions (C1)–(C5) hold. Then

$$\sup_{0 < u < 1} \|\tilde{\mu}(u) - \mu(u)\| = O_p(C_n), \quad (A.3)$$

$$\sup_{0 < u < 1} \|\tilde{\theta}(u) - \theta(u)\| = O_p(C_n),$$

where  $C_n = \{\log(1/h)/nh\}^{1/2} + h^2$ .

*Proof.* Let  $S_{nl} = \sum_{i=1}^n X_i X_i^T ((U_i - u)/h)^l K_h(U_i - u)$  and let  $l = 0, 1, 2$ . Similar to the proof of Lemma 7.2 in [7], a simple calculation yields

$$E(S_{nl}) = nf(u)\Psi(u) \int_0^1 s^l K(s) ds + o(1), \quad l = 0, 1, 2. \quad (A.4)$$

Note that  $D_u^T \Omega_u D_u = \begin{pmatrix} S_{n0} & S_{n1} \\ S_{n1} & S_{n2} \end{pmatrix}$ , by Lemma 3, we obtain

$$D_u^T \Omega_u D_u = nf(u)\Psi(u) \otimes \left( \begin{matrix} 1 & 0 \\ 0 & \int_0^1 t^2 K(t) dt \end{matrix} \right) \{1 + O_p(C_n)\}, \quad (A.5)$$

TABLE 1: The average lengths and coverage probabilities of the 95% confidence intervals for  $\beta$  based on the AEL method and the NA method.

$n$	CR	Confidence interval		Length		Coverage	
		AEL	NA	AEL	NA	AEL	NA
50	0.1	(0.924, 1.247)	(0.828, 1.344)	0.323	0.516	0.914	0.916
	0.3	(0.915, 1.259)	(0.825, 1.347)	0.344	0.522	0.896	0.894
	0.45	(0.906, 1.264)	(0.824, 1.350)	0.358	0.526	0.863	0.864
100	0.1	(0.927, 1.220)	(0.913, 1.270)	0.293	0.357	0.935	0.937
	0.3	(0.922, 1.223)	(0.907, 1.273)	0.301	0.366	0.931	0.930
	0.45	(0.916, 1.227)	(0.906, 1.276)	0.311	0.370	0.928	0.928
150	0.1	(0.960, 1.161)	(0.942, 1.239)	0.201	0.297	0.948	0.948
	0.3	(0.952, 1.199)	(0.939, 1.237)	0.247	0.298	0.945	0.943
	0.45	(0.949, 1.202)	(0.933, 1.236)	0.253	0.303	0.943	0.942

TABLE 2: Application to CGD data. The maximum empirical likelihood estimators (MELE) and 95% confidence intervals (CI) for the parametric components based on the AEL method.

	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$
MELE	4.470	-0.288	0.013	0.004	-0.007
CI	(4.210, 4.680)	(-0.473, -0.158)	(0.002, 0.021)	(0.002, 0.005)	(-0.012, -0.004)
	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	
MELE	0.398	-0.097	0.305	-0.021	
CI	(0.268, 0.501)	(-0.267, 0.033)	(0.178, 0.412)	(-0.128, 0.045)	

uniformly for  $u \in (0, 1)$ , where  $\otimes$  is the Kronecker product. Using the same argument, we have

$$D_u^T \Omega_u Z = n f(u) \Phi(u) \otimes (1, 0)^T \{1 + O_p(C_n)\}, \quad (\text{A.6})$$

uniformly for  $u \in (0, 1)$ . Combining (A.5) and (A.6) yields

$$\begin{aligned} \tilde{\mu}(u) &= (I_p, 0_p) (D_u^T \Omega_u D_u)^{-1} D_u^T \Omega_u Z \\ &= \mu(u) \{1 + O_p(C_n)\}, \end{aligned} \quad (\text{A.7})$$

uniformly for  $u \in (0, 1)$ .

Invoking  $E(Y_{iG}^* | X_i, Z_i, U_i) = E(Y_i | X_i, Z_i, U_i)$ , with the similar argument, we can prove  $\sup_{0 < u < 1} \|\tilde{\theta}(u) - \theta(u)\| = O_p(C_n)$ . This completes the proof of Lemma 4.  $\square$

**Lemma 5.** Suppose that conditions (C1)–(C5) hold. If  $\beta_0$  is the true parameter, then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \hat{\eta}_i(\beta_0) \xrightarrow{\mathcal{L}} N(0, \Sigma(\beta_0)), \quad (\text{A.8})$$

where  $\Sigma(\beta_0)$  is defined in Theorem 1.

*Proof.* By (8), we have

$$\begin{aligned} \hat{\eta}_i(\beta_0) &= \check{Z}_i (\check{Y}_{iG}^* - \check{Z}_i^T \beta_0) + \check{Z}_i (\check{Y}_{iG_n}^* - \check{Y}_{iG}^*) \\ &\equiv J_{1i}(\beta_0) + J_{2i}(\beta_0). \end{aligned} \quad (\text{A.9})$$

A simple calculation yields

$$\begin{aligned} &\frac{1}{\sqrt{n}} \sum_{i=1}^n J_{1i}(\beta_0) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n [Z_i - \mu(U_i)^T X_i] [Y_{iG}^* - X_i^T \theta(U_i) - Z_i^T \beta_0] \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n [Z_i - \mu(U_i)^T X_i] X_i^T [\theta(U_i) - \tilde{\theta}(U_i)] \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n [\mu(U_i) - \tilde{\mu}(U_i)]^T X_i [Y_{iG}^* - X_i^T \theta(U_i) - Z_i^T \beta_0] \\ &\quad + \frac{1}{\sqrt{n}} \sum_{i=1}^n [\mu(U_i) - \tilde{\mu}(U_i)]^T X_i X_i^T [\theta(U_i) - \tilde{\theta}(U_i)] \\ &\equiv A_1 + A_2 + A_3 + A_4. \end{aligned} \quad (\text{A.10})$$

By the Central Limits Theorem, it is easy to prove  $A_1 \xrightarrow{\mathcal{L}} N(0, \Sigma_0(\beta_0))$ . Next, we prove  $A_\nu = o_p(1)$ ,  $\nu = 2, 3, 4$ . Using Abel inequality, invoking  $E\{(Z_i - \mu(U_i)^T X_i) X_i^T\} = 0$ , and using Lemma 4, we can prove that

$$\begin{aligned} \|A_2\| &\leq \frac{C}{\sqrt{n}} \sup_{1 \leq i \leq n} \|\tilde{\theta}(U_i) - \theta(U_i)\| \\ &\quad \times \max_{1 \leq k \leq n} \left\| \sum_{i=1}^k [Z_i - \mu(U_i)^T X_i] X_i^T \right\| \\ &= \frac{C}{\sqrt{n}} O_p(C_n) O_p(\sqrt{n} \log n), \end{aligned} \quad (\text{A.11})$$

that is,  $A_2 = o_p(1)$ . Invoking  $E(Y_{iG}^* | X_i, Z_i, U_i) = E(Y_i | X_i, Z_i, U_i)$ , and with a similar argument, we can prove  $A_3 = o_p(1)$ . In addition, by Lemma 4, we have  $\|A_4\| \leq O_p(\sqrt{n}C_n^2) = o_p(1)$ . Hence, we get that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n J_{1i}(\beta_0) \xrightarrow{\mathcal{L}} N(0, \widehat{\Sigma}_0(\beta_0)). \quad (\text{A.12})$$

By Wang and Zheng [1], we have

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^n J_{2i}(\beta_0) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{[Z_i - \mu(U_i)]^T X_i}{1 - G(Y_i^*)} \frac{Y_i^* \delta_i G_n(Y_i^*) - G(Y_i^*)}{1 - G(Y_i^*)} \\ &+ o_p(1). \end{aligned} \quad (\text{A.13})$$

Then, using the similar argument to Yang et al. [3], we can prove that

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{i=1}^n J_{2i}(\beta_0) \xrightarrow{\mathcal{L}} N(0, \Sigma_1), \\ & \frac{1}{n} \sum_{i=1}^n E\{J_{1i} J_{2i}^T\} \xrightarrow{P} \Sigma_1. \end{aligned} \quad (\text{A.14})$$

Together with (A.12), we complete the proof of Lemma 5.  $\square$

*Proof of Theorem 1.* From the proof of Lemma 5, it is easy to show that

$$\max_{1 \leq i \leq n} \|\widehat{\eta}_i(\beta_0)\| = o_p(n^{1/2}). \quad (\text{A.15})$$

Then, using the arguments similar to Owen [9], we can obtain  $\|\lambda\| = O_p(n^{-1/2})$ . Together with Lemma 5, applying the Taylor expansion to (10), we get that

$$\widehat{R}(\beta_0) = 2 \sum_{i=1}^n \left\{ \lambda^T \widehat{\eta}_i(\beta_0) - \frac{[\lambda^T \widehat{\eta}_i(\beta_0)]^2}{2} \right\} + o_p(1). \quad (\text{A.16})$$

By (11), it follows that

$$\begin{aligned} 0 &= \sum_{i=1}^n \frac{\widehat{\eta}_i(\beta_0)}{1 + \lambda^T \widehat{\eta}_i(\beta_0)} \\ &= \sum_{i=1}^n \widehat{\eta}_i(\beta_0) - \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \widehat{\eta}_i^T(\beta_0) \lambda \\ &+ \sum_{i=1}^n \frac{\widehat{\eta}_i(\beta_0) [\lambda^T \widehat{\eta}_i(\beta_0)]^2}{1 + \lambda^T \widehat{\eta}_i(\beta_0)}. \end{aligned} \quad (\text{A.17})$$

Then, it is easy to show that

$$\sum_{i=1}^n [\lambda^T \widehat{\eta}_i(\beta_0)]^2 = \sum_{i=1}^n \lambda^T \widehat{\eta}_i(\beta_0) + o_p(1), \quad (\text{A.18})$$

$$\lambda = \left\{ \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \widehat{\eta}_i^T(\beta_0) \right\}^{-1} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) + o_p(n^{-1/2}).$$

Invoking (A.16)–(A.18), by some algebra calculations, we have

$$\widehat{R}(\beta_0) = \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \right\}^T \widehat{\Sigma}_0^{-1}(\beta_0) \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \right\}, \quad (\text{A.19})$$

where  $\widehat{\Sigma}_0(\beta_0) = n^{-1} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \widehat{\eta}_i^T(\beta_0)$ . Invoking the proof of Lemma 5, we can obtain that  $\widehat{\Sigma}_0(\beta_0) \xrightarrow{P} \Sigma_0(\beta_0)$ . Hence, we get

$$\begin{aligned} \widehat{R}(\beta_0) &= \left( \Sigma^{-1/2}(\beta_0) \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \right)^T \Gamma(\beta_0) \\ &\times \left( \Sigma^{-1/2}(\beta_0) \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \right) + o_p(1), \end{aligned} \quad (\text{A.20})$$

where  $\Gamma(\beta_0) = \Sigma^{1/2}(\beta_0) \Sigma_0^{-1}(\beta_0) \Sigma^{1/2}(\beta_0)$ . Let  $w_1, \dots, w_q$  be the eigenvalues of  $\Sigma_0^{-1}(\beta_0) \Sigma(\beta_0)$ , and let  $D = \text{diag}(w_1, \dots, w_q)$ . Note that  $\Sigma_0^{-1}(\beta_0) \Sigma(\beta_0)$  has the same eigenvalues as  $\Gamma(\beta_0)$ . Hence, there exists orthogonal matrix  $Q$  such that  $Q^T D Q = \Gamma(\beta_0)$ . Then, with a simple calculation we get that

$$\begin{aligned} \widehat{R}(\beta_0) &= \left( Q \Sigma^{-1/2}(\beta_0) \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \right)^T D \\ &\times \left( Q \Sigma^{-1/2}(\beta_0) \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \right) + o_p(1). \end{aligned} \quad (\text{A.21})$$

Note that  $Q$  is an orthogonal matrix, and by Lemma 5, Theorem 1 can be proved.  $\square$

*Proof of Theorem 2.* Let

$$\Xi(\beta_0) = \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \right\}^T \widehat{\Sigma}^{-1}(\beta_0) \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{\eta}_i(\beta_0) \right\}. \quad (\text{A.22})$$

It is easy to show that  $\widehat{R}_{\text{ad}}(\beta_0) = \Xi(\beta_0) + o_p(1)$ . Then, using similar arguments to Yang et al. [3], it can be proved that

$$\widehat{\Sigma}(\beta_0) \xrightarrow{P} \Sigma(\beta_0). \quad (\text{A.23})$$

Together with Lemma 5, we can prove Theorem 2.  $\square$

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (11101119 and 11126332), the National Social Science Foundation of China (11CTJ004), the Natural Science Foundation of Guangxi (2010GXNSFB013051), and the Philosophy and Social Sciences Foundation of Guangxi (11FTJ002).

## References

- [1] Q. Wang and Z. Zheng, "Asymptotic properties for the semi-parametric regression model with randomly censored data," *Science in China A*, vol. 40, no. 9, pp. 945–957, 1997.
- [2] 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.
- [3] Y. P. Yang, L. G. Xue, and W. H. Cheng, "An empirical likelihood method in a partially linear single-index model with right censored data," *Acta Mathematica Sinica*, vol. 28, no. 5, pp. 1041–1060, 2012.
- [4] X. Lu and T. L. Cheng, "Randomly censored partially linear single-index models," *Journal of Multivariate Analysis*, vol. 98, no. 10, pp. 1895–1922, 2007.
- [5] Q. Wang and L. Xue, "Statistical inference in partially-varying-coefficient single-index model," *Journal of Multivariate Analysis*, vol. 102, no. 1, pp. 1–19, 2011.
- [6] Z. Huang, "Empirical likelihood-based inference in varying-coefficient single-index models," *Journal of the Korean Statistical Society*, vol. 40, no. 2, pp. 205–215, 2011.
- [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] 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.
- [9] A. B. Owen, "Empirical likelihood ratio confidence regions," *Annals of Statistics*, vol. 18, no. 1, pp. 90–120, 1990.
- [10] L. Xue, "Empirical likelihood local polynomial regression analysis of clustered data," *Scandinavian Journal of Statistics*, vol. 37, no. 4, pp. 644–663, 2010.
- [11] Y. P. Yang, L. G. Xue, and W. H. Cheng, "Empirical likelihood for a partially linear model with covariate data missing at random," *Journal of Statistical Planning and Inference*, vol. 139, no. 12, pp. 4143–4153, 2009.
- [12] X. Luo, Y. Li, Y. Ma, and Y. Zhou, "Varying-coefficient partially linear models with censored data," *Acta Mathematica Scientia*, vol. 26, pp. 79–92, 2010.
- [13] T. R. Fleming and D. P. Harrington, *Counting Processes and Survival Analysis*, John Wiley and Sons, New York, NY, USA, 2005.
- [14] R. Jiang and W. M. Qian, "Generalized likelihood ratio tests for varying-coefficient models with censored data," *Open Journal of Statistics*, vol. 1, no. 1, pp. 19–23, 2011.
- [15] 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.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

