

Asymptotic Normality in Varying-Coefficient Errors-in-Variables Models with Missing Response Variables

YaFeng Xia

College of Science Lanzhou University of Technology
Lanzhou, China
e-mail: gsxyf01@163.com

Yu Bai

College of Science Lanzhou University of Technology
Lanzhou, China
e-mail: 455047015@qq.com

Abstract—This article is concerned with the estimation of a varying-coefficient regression model when the explanatory variables are measured with additive errors and the response variable is sometimes missing. Estimated coefficient function in complete observational data and interpolation data respectively, and all the proposed estimators for the coefficient function are proved to be asymptotic normality. Finally, a simulation study is conducted to compare the proposed estimators.

Keywords-missing data; local linear least squares method; varying-coefficient models; errors-in-variables

I. INTRODUCTION

The form of varying coefficient models is:

$$Y = X^T \beta(U) + \varepsilon \quad (1.1)$$

where Y is the response variable, and X is U the explanatory variables, we assume that $\beta(\cdot) = (\beta_1(\cdot), \dots, \beta_p(\cdot))^T$ is a p -dimensional function of unknown coefficients, U is a single variable, ε is a random statistical error with $E[\varepsilon|X, U] = 0$ and $Var[\varepsilon|X, U] = \sigma^2$.

Since *Hastie* and *Tibshirani*(1993) proposed the varying coefficient regression models, it had been developing rapidly and applied widely in many fields because of their flexibility and adaptability, such as, *Fan*, and *Zhang*(1999), *Chiang*, *Rice* and *Wu*(2001), *Huang*, *Wu* and *Zhou*(2002), *Hoover*, *Wu*, and *Yang*(1998), and *Wu*, *Chiang*, and *Hoover*(1998). But most of their attentions and works focused on the case where all variables with no missing data or measurement error, so it will be a meaningful thing for studying varying-coefficient errors-in-variables models with missing response variables.

Based on the profile least square technology and the correction technique, *Wei*(2010) developed some approaches of estimating β and $\alpha(x)$ for partially linear models in complete observational data, interpolation data, and the surrogate data respectively, finally, proved the asymptotic normality of β and $\alpha(x)$. *Wei*(2011) studied coefficient function's estimation for varying-coefficient errors-in-variables models with missing response variables, and proved its asymptotic normality.

les models with missing response variables, and proved its asymptotic normality.

In this paper, based on *Wei*(2010,2011), suppose we obtain a random sample of incomplete data $\{(X_i, Y_i, \delta_i), 1 \leq i \leq n\}$ from model(1.1), where $\delta_i = 0$ if Y_i is missing and $\delta_i = 1$ otherwise. Namely, the missing data mechanism is $pr(\delta_i = 1 | Y_i, X_i, U_i) = pr(\delta_i = 1 | X_i, U_i) = \pi(X_i, U_i)$. Throughout this paper, we assume that Y may be missing and X is observed completely, the explanatory variables have measurement errors, namely, $V = X + \xi$, where $(X^T, U, \delta)^T$ and ξ are independent, $Var[\xi|X, U] = \Sigma_\xi$ and $E[\xi|X, U] = 0$. We assumed Σ_ξ is known, if the covariance matrix is unknown, it can be estimated using repeated measures data X . Next, by using local linear least square method, the correction technique and the interpolation method, we estimated coefficient functions in the complete data and interpolation data, and the estimated asymptotic normality are proved.

Next, by using local linear least square method, the correction technique and the interpolation method, we estimated coefficient functions in the complete data and interpolation data, and the estimated asymptotic normality are proved.

II. THE ESTIMATION METHODS

A. Complete Observation data Estimation Method

For the regression model with missing response variables, a simple and direct research method is complete data method(There is no missing part of the complete observation response variables for statistical inference), that is observation data corresponding to $\delta_i = 1$, there are

$$\delta_i Y_i = \delta_i X_i^T \beta(U_i) + \delta_i \varepsilon_i, i = 1, \dots, n \quad (2.1)$$

where $(Y_i, X_i, U_i)_{i=1}^n$ is a independent identically distributed sample from model (1.1), we use a local linear least square method to estimate the coefficient function $\beta_j(u) (j=1, \dots, p)$,

for every $\beta_j(u)$, it has a continuous second derivative, then, $\beta_j(u)$ can be approximate locally at $\beta_j(u_0)$ by a linear function:

$$\beta_j(u) \approx \beta_j(u_0) + \beta_j'(u_0)(u - u_0) \equiv a_j + b_j(u - u_0), (j=1, \dots, p)$$

$\beta_j(u_0)$, $\beta_j'(u_0)$ can be estimated by minimizing

$$\sum_{i=1}^n [Y_i - \sum_{j=1}^p \{a_j + b_j(U_i - u_0)\} X_{ij}]^2 K_h(U_i - u_0) \delta_i \quad (2.2)$$

where, $K(\cdot)$ is a kernel function, h is a bandwidth, $K_h(\cdot) = K(\cdot/h)/h$.

Let $\theta(u_0)=(\beta_j(u_0)^T, h\beta'_j(u_0)^T)^T, \eta_i=(X_i^T, \frac{U_i-u_0}{h} X_i^T)^T$, we can get the following equation on the basis of (2.2)

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \phi_{11} \{ \theta(u_0); Y_i, X_i, U_i \} \\ &= \frac{1}{n} \sum_{i=1}^n \delta_i K_h(U_i-u_0) (\eta_i \eta_i^T \theta(u_0) - \eta_i Y_i) \end{aligned} \quad (2.3)$$

because X_i is not accurate observation, if ignored the error of measurement, we replace X_i with V_i , it is easy to prove the estimates obtained is not consistent, therefore we remember:

$$E(V_i V_i^T | X_i, U_i, Y_i) = E(X_i X_i^T) + \Sigma_\xi, E(V_i Y_i^T | X_i, U_i, Y_i) = E(X_i Y_i)$$

Based on the correction technique, then we can get the following correction estimation equation:

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \phi_{11}^* \{ \theta(u_0); Y_i, X_i, U_i \} \\ &= \frac{1}{n} \sum_{i=1}^n \delta_i K_h(U_i-u_0) \{ (\eta_i^* \eta_i^{*T} - \gamma_i) \theta(u_0) - \eta_i^* Y_i \} \end{aligned} \quad (2.4)$$

where, η_i^* , X_i were, respectively, replaced with η_i , V_i .

$$\gamma_i = \Sigma_\xi \otimes \begin{pmatrix} 1 & (U_i-u_0)/h \\ (U_i-u_0)/h & [(U_i-u_0)/h]^2 \end{pmatrix}$$

The solution to (2.4), we have the following estimates:

$$\begin{aligned} \hat{\theta}(u_0) &= \left\{ \frac{1}{n} \sum_{i=1}^n \delta_i K_h(U_i-u_0) (\eta_i^* \eta_i^{*T} - \gamma_i) \right\}^{-1} \left\{ \frac{1}{n} \sum_{i=1}^n \delta_i K_h(U_i-u_0) \eta_i^* Y_i \right\} \\ &= \{ D_{u_0}^T W_{u_0}^\delta D_{u_0} - \Omega_{u_0}^\delta \}^{-1} D_{u_0}^T W_{u_0}^\delta Y \end{aligned} \quad (2.5)$$

where, $W_{u_0}^\delta = \text{diag} \{ \delta_1 K_h(U_1-u_0), \dots, \delta_n K_h(U_n-u_0) \}$,

$$D_{u_0} = \begin{bmatrix} V_1^T & \frac{U_1-u_0}{h} V_1^T \\ V_2^T & \frac{U_2-u_0}{h} V_2^T \\ \vdots & \vdots \\ V_n^T & \frac{U_n-u_0}{h} V_n^T \end{bmatrix}, Y = (Y_1, Y_2, \dots, Y_n)^T$$

$$\Omega_{u_0}^\delta = \sum_{i=1}^n \Sigma_\xi \otimes \begin{pmatrix} 1 & (U_i-u_0)/h \\ (U_i-u_0)/h & [(U_i-u_0)/h]^2 \end{pmatrix} K_h(U_i-u_0) \delta_i$$

According to (2.5), the coefficient function is estimated as follows

$$\hat{\beta}(u) = (I_p \ O_p) \{ D_u^T W_u^\delta D_u - \Omega_u^\delta \}^{-1} D_u^T W_u^\delta Y \quad (2.6)$$

B. Interpolation Estimation Method

In order to make full use of data information, we use the interpolation method of *chu* and *cheng*(1995). The idea of the method is the use of single point interpolation technique for filling missing variables, then all data involved in the back of the inference process. Specifically, by using the estimation of $\hat{\beta}(u)$ obtained in the complete observation

data of section 2.1, if the X_i is observed accurately, we have observed data $(Y_i^0, X_i, U_i)_{i=1}^n$, where $Y_i^0 = \delta_i Y_i + (1-\delta_i) X_i^T$

$\hat{\beta}(u_i)$, but because of the exact value of X_i does not exist, we can't obtain the exact value of Y_i^0 and we can get $Y_i^* = \delta_i Y_i + (1-\delta_i) V_i^T \hat{\beta}(u_i)$. Then we can get the varying coefficient model with missing response variables and measurement errors of the explanatory variable on the basis of $(Y_i^*, X_i, V_i, U_i)_{i=1}^n$

$$\begin{cases} Y_i^0 = X_i^T \beta(U_i) + e_i \\ V_i = X_i + \xi_i \\ Y_i^* = Y_i^0 + (1-\delta_i) \xi_i^T \hat{\beta}(u_i) \end{cases} \quad i=1, \dots, n \quad (2.7)$$

where $e_i = Y_i^0 - X_i + \varepsilon_i$ is the model error.

Similar to (2.2), we can through the minimization

$$\sum_{i=1}^n [Y_i^* - \sum_{j=1}^p \{ a_j + b_j(U_i-u_0) \} X_{ij}]^2 K_h(U_i-u_0) \quad (2.8)$$

Further to estimate the following equation

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \phi_{12} \{ \theta(u_0); Y_i^*, X_i, U_i \} \\ &= \frac{1}{n} \sum_{i=1}^n K_h(U_i-u_0) (\eta_i \eta_i^T \theta(u_0) - \eta_i Y_i^*) \end{aligned} \quad (2.9)$$

Among them, because of the errors, we remember:

$$E(V_i V_i^T | X_i, U_i, Y_i^*) = E(X_i X_i^T) + \Sigma_\xi, E(V_i Y_i^{*T} | X_i, U_i, Y_i^*) = E(X_i Y_i^{*T})$$

Based on the correction technique, then we can get the following correction estimation equation

$$\begin{aligned} 0 &= \frac{1}{n} \sum_{i=1}^n \phi_{12}^* \{ \theta(u_0); Y_i^*, X_i, U_i \} \\ &= \frac{1}{n} \sum_{i=1}^n K_h(U_i-u_0) \{ (\eta_i^{**} \eta_i^{**T} - \gamma_i) \theta(u_0) - \eta_i^{**} (Y_i^* - (1-\delta_i) \Sigma_\xi \hat{\beta}(u_0)) \} \end{aligned} \quad (2.10)$$

where, η_i^{**} , X_i were, respectively, replaced with η_i , V_i .

The solution to (2.10), we have the following estimates:

$$\begin{aligned} \tilde{\theta}(u_0) &= \left\{ \frac{1}{n} \sum_{i=1}^n K_h(U_i-u_0) (\eta_i^{**} \eta_i^{**T} - \gamma_i) \right\}^{-1} \left\{ \frac{1}{n} \sum_{i=1}^n K_h(U_i-u_0) \eta_i^{**} \right. \\ &\quad \left. (Y_i^* - (1-\delta_i) \Sigma_\xi \hat{\beta}(u_0)) \right\} = \{ D_{u_0}^T W_{u_0} D_{u_0} - \Omega_{u_0} \}^{-1} D_{u_0}^T W_{u_0} (Y^* - C_{u_0}) \end{aligned} \quad (2.11)$$

Where, $W_{u_0} = \text{diag} \{ K_h(U_1-u_0), \dots, K_h(U_n-u_0) \}$,

$$\Omega_{u_0} = \sum_{i=1}^n \Sigma_\xi \otimes \begin{pmatrix} 1 & (U_i-u_0)/h \\ (U_i-u_0)/h & [(U_i-u_0)/h]^2 \end{pmatrix} K_h(U_i-u_0),$$

$$C_{u_0} = \left((1-\delta_1) \Sigma_\xi \hat{\beta}(u_0), \dots, (1-\delta_n) \Sigma_\xi \hat{\beta}(u_0) \right)^T,$$

$$Y^* = (Y_1^*, Y_2^*, \dots, Y_n^*)^T$$

According to (2.11), the coefficient function is estimated as follows

$$\tilde{\beta}(u) = (I_p \ O_p) \{D_u^T W_u D_u - \Omega_u\}^{-1} D_u^T W_u (Y^* - C_u)$$

III. MAIN RESULTS

The following assumptions we need:

(A1) $(X_i^T, U_i, \varepsilon_i, \xi_i^T)^T$ are independent and identically

distributed random vectors, $(X_i^T, U_i)^T$ and $(\varepsilon_i, \xi_i^T)^T$ are independent, ε_i and ξ_i are independent.

(A2) the random variable U has a bounded support Π , its density function $f(\cdot)$ is Lipschitz continuous, and not to 0 on the support.

(A3) the $p \times p$ matrix $E[\pi(X, U) X X^T | U]$ is nonsingular for any $U \in \Pi$.

(A4) There is an $s > 2$ such that $E\|X\|^{2s} < \infty$ and $E\|\xi\|^{2s} < \infty$ and for some $k < 2 - s^{-1}$ such that $n^{2k-1} h \rightarrow \infty$ as $n \rightarrow \infty$.

(A5) $\beta(\cdot)$ is a second order continuous derivatives in $U \in \Pi$.

(A6) $K(\cdot)$ is a function with compact support symmetric density function, and the bandwidth h satisfies $nh^8 \rightarrow 0$ and $nh^2 / (\log n)^2 \rightarrow \infty$ and, as $n \rightarrow \infty$.

Let

$$\mu_i(K) = \int_0^\infty t^i K(t) dt, \quad \beta''(u) = [\beta_1''(u), \dots, \beta_p''(u)]^T,$$

$$v_i(K) = \int_0^\infty t^i K^2(t) dt, \quad \beta_j''(u) = \partial^2 \beta_j(u) / \partial u^2,$$

then we obtain the following theorem properties:

Theorem 1. Assuming the above assumptions, we can get the following conclusion

$$\sqrt{nh} \left\{ \hat{\beta}(u) - \beta(u) - \frac{h^2}{2} \mu_2(K) \beta''(u) + o_P(h^2) \right\} \xrightarrow{d} N \left(0, \frac{v_0(K)}{f(u)} \Gamma_\delta(u)^{-1} \Omega_\delta(u) \Gamma_\delta(u)^{-1} \right)$$

Where

$$\Omega_\delta(u) = \sigma^2 \Gamma_\delta(u) + \sigma^2 \Sigma_\xi E(\delta_1 | U_1 = u) + E[\delta_1 \eta \beta(u) \beta^T(u) \eta_1^T | U_1 = u]$$

$$\Gamma_\delta(u) = E \left\{ \delta_1 X_1 X_1^T | U_1 = u \right\}, \quad \eta_i = \Sigma_\xi - \xi_i \xi_i^T - X_i \xi_i^T$$

In order to make statistical inference with theorem 1,

$$\text{let } \hat{\Gamma}_\delta(u) = n^{-1} \sum_{i=1}^n \delta_i K_h(U_i - u) [V_i V_i^T - \Sigma_\xi]$$

$$\hat{\Omega}_\delta(u) = n^{-1} \sum_{i=1}^n \delta_i K_h^2(U_i - u) \{V_i [Y_i - V_i^T \hat{\beta}(u_i)] + \Sigma_\xi \hat{\beta}(u)\} \otimes 2$$

And it is easy to prove that $\hat{\Gamma}_\delta(u)^{-1} \hat{\Omega}_\delta(u) \hat{\Gamma}_\delta(u)^{-1}$ is a

consistent estimate of $\frac{v_0(K)}{f(u)} \Gamma_\delta(u)^{-1} \Omega_\delta(u) \Gamma_\delta(u)^{-1}$.

Theorem 2. Assuming the above assumptions, we can get the following conclusion

$$\sqrt{nh} \left\{ \tilde{\beta}(u) - \beta(u) - \frac{h^2}{2} \mu_2(K) \beta''(u) + o_P(h^2) \right\} \xrightarrow{d} N \left(0, \frac{v_0(K)}{f(u)} \Gamma(u)^{-1} \Omega(u) \Gamma(u)^{-1} \right)$$

where $\Omega(u) = \sigma^2 \Gamma(u) + \sigma^2 \Sigma_\xi + E[\eta \beta(u) \beta^T(u) \eta_1^T | U_1 = u]$,

$$\Gamma(u) = E \left\{ X_1 X_1^T | U_1 = u \right\}, \quad \eta_i = \Sigma_\xi - \xi_i \xi_i^T - X_i \xi_i^T,$$

IV. NUMETRICAL SIMULATION

In this section, by adopting stochastic simulation method, we will discuss the fitting degree of the estimators in the case of complete data and interpolation data and the real function. we consider the following model

$$y_i = x_i^T \beta(u_i) + \varepsilon_i, \quad v_i = x_i + e_i, \quad x = 1, 2, \dots, n,$$

where $x_i \rightarrow N(0, 1)$, $u_i \rightarrow U(0, 1)$, $e_i \rightarrow N(0, 0.1^2)$, $\beta(u_i) = \cos(6\pi u_i)$

the kernel function is $K(x) = \frac{15}{16} (1 - x^2)^2 I_{|x| \leq 1}$. In order to studying the effect of model error distribution to the final result, we consider the following two kinds of distribution of ε_i , that is $\varepsilon_i \rightarrow N(0, 0.1)$ and $\varepsilon_i \rightarrow U(-0.1, 0.1)$. If $X_i \leq 0.8$, $\delta_i = 1$, else $\delta_i = 0$, the independent random missing data is produced like this.

Our simulation results are as follows (To name only a few of the main results). We are to estimate the 100 time point on $[0, 1]$. In the following figure, $\beta(u)$ is showed by blue line, $\hat{\beta}(u)$ and $\tilde{\beta}(u)$ is showed by '*'.

During the simulation, we consider the following situations

(1) figure 1, 2: the sample size $n=1000$, the bandwidth $h=1/50$, $\varepsilon_i \rightarrow N(0, 0.1)$

(2) figure 3, 4: the sample size $n=1000$, the bandwidth $h=1/100$, $\varepsilon_i \rightarrow N(0, 0.1)$

(3) figure 5, 6: the sample size $n=1000$, the bandwidth $h=1/100$, $\varepsilon_i \rightarrow U(-0.1, 0.1)$

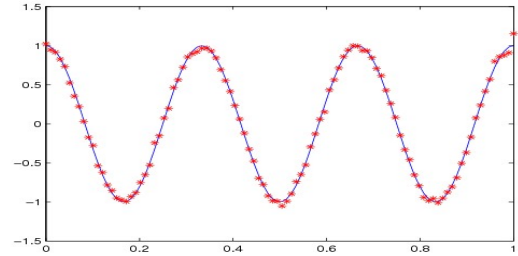


Figure1 $\beta(u)$ and $\hat{\beta}(u)$

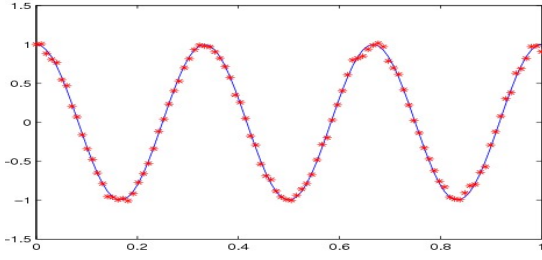


Figure2 $\beta(u)$ and $\hat{\beta}(u)$

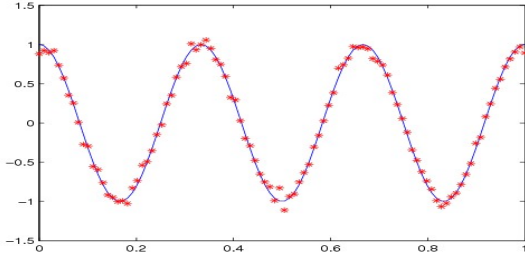


Figure3 $\beta(u)$ and $\hat{\beta}(u)$

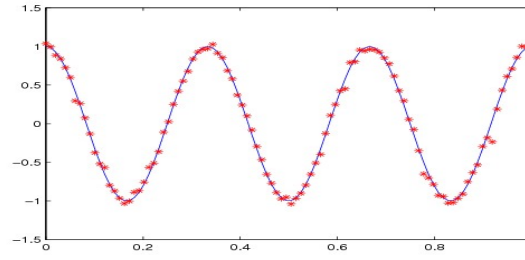


Figure4 $\beta(u)$ and $\hat{\beta}(u)$

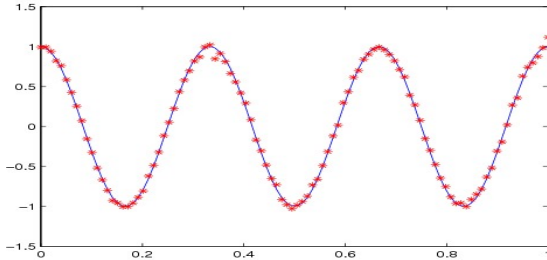


Figure5 $\beta(u)$ and $\hat{\beta}(u)$

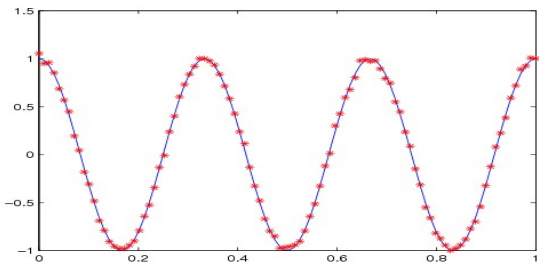


Figure6 $\beta(u)$ and $\hat{\beta}(u)$

Generally, the simulation results show that the method we used to estimate is good. From our simulation result shows, when the sample size from 1000 generally can get good effect, due to the total simulation effect is good, the two window width here seems to be little difference. At the same time also can see the impact on the simulation results of the model error distribution also is not very big. And we found the interpolation estimates $\hat{\beta}(u)$ is closer to the real function of value under the relatively complete data $\hat{\beta}(u)$.

V. PROOF OF THEOREM

In order to prove the main results, we first introduce some lemmas and notations, note

$$\Delta = \text{diag}\{\delta_1, \dots, \delta_n\}, M = \left[X_1^T \beta(u_1), \dots, X_p^T \beta(u_p) \right]^T$$

$$v_i(K) = \int_0^\infty t^i K^2(t) dt, \mu_i(K) = \int_0^\infty t^i K(t) dt, c_n = h^2 + \left\{ \frac{\log(1/h)}{nh} \right\}^{1/2}$$

Lemma 1. Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d random vectors, where $Y_i, i=1, 2, \dots, n$ is a random variable, Further, assume that

$E|y|^s < \infty$ and $\sup_x \int |y|^s f(x, y) dy < \infty$, f denote the joint density of (X, Y) . Let K is a bounded positive function with a bounded support, satisfying a Lipschitz condition. Then

$$\sup_x \left| \frac{1}{n} \sum_{i=1}^n [K_h(X_i - X) Y_i - E(K_h(X_i - X) Y_i)] \right| = O_p \left(\left\{ \frac{\log(1/h)}{nh} \right\}^{1/2} \right)$$

where there are $n^{2\varepsilon-1} h \rightarrow \infty$ as $\varepsilon < 1-s^{-1}$.

Proof. This lemma can be obtained by Mack and Silverman (1982).

Lemma 2. Under the conditions in Sec.5, we have

$$\frac{D_u^T W_u^\delta D_u - \Omega_u^\delta}{n} \xrightarrow{p} f(u) \Gamma(u) \otimes \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 \end{pmatrix} + o_p(1)$$

Proof. The proof of lemma similar to Cai et al.(2000) proof of theorem 1, we ignore the detailed proof.

Proof of Theorems 1. Because in the field of u , $\beta_j(\cdot)$ is smooth, by Taylor expansion, we can obtain:

$$X_i^T \beta(U_i) = X_i^T \beta(u) + X_i^T h \beta'(u) \frac{U_i - u}{h} + \frac{h^2}{2} X_i^T \beta''(u) \left(\frac{U_i - u}{h} \right)^2 + o_p(h^2)$$

By the definition of $\hat{\beta}(u)$, we have

$$\begin{aligned} \tilde{\beta}(u) - \beta(u) &= (I_p O_p) \{ D_u^T W_u^\delta D_u - \Omega_u^\delta \}^{-1} D_u^T W_u^\delta Y - \beta(u) \\ &= (I_p O_p) \{ D_u^T W_u^\delta D_u - \Omega_u^\delta \}^{-1} D_u^T W_u^\delta A \frac{h^2 \beta''(u)}{2} \\ &\quad + (I_p O_p) \{ D_u^T W_u^\delta D_u - \Omega_u^\delta \}^{-1} D_u^T W_u^\delta I_n o_p(h^2) \\ &\quad + (I_p O_p) \{ D_u^T W_u^\delta D_u - \Omega_u^\delta \}^{-1} D_u^T W_u^\delta \varepsilon \\ &\quad + (I_p O_p) \{ D_u^T W_u^\delta D_u - \Omega_u^\delta \}^{-1} (\Omega_u^\delta - D_u^T W_u^\delta B) \theta(u) \end{aligned}$$

$$= I_1 + I_2 + I_3 + I_4$$

where

$$A = \begin{bmatrix} X_1^T \left(\frac{U_1-u}{h} \right)^2 \\ X_2^T \left(\frac{U_2-u}{h} \right)^2 \\ \vdots \\ X_n^T \left(\frac{U_n-u}{h} \right)^2 \end{bmatrix}, B = \begin{bmatrix} \xi_1^T & \frac{U_1-u}{h} \xi_1^T \\ \xi_2^T & \frac{U_2-u}{h} \xi_2^T \\ \vdots & \vdots \\ \xi_n^T & \frac{U_n-u}{h} \xi_n^T \end{bmatrix}$$

By lemma 1 and lemma 2, it is easy to prove

$$I_1 = \frac{h^2}{2} \mu_2(K) \beta''(u) + o_p(1), I_2 = o_p(h^2)$$

$$\text{and } \sqrt{nh} \left\{ \hat{\beta}(u) - \beta(u) - \frac{h^2}{2} \mu_2(K) \beta''(u) + o_p(h^2) \right\} \\ = f(u)^{-1} \Gamma(u)^{-1} \sqrt{nh} \frac{1}{n} \sum_{i=1}^n K_h(U_i-u) \{V_i \varepsilon_i + \eta_i \beta(u)\} + o_p(1)$$

Therefore, by the central limit theorem can be

$$\sqrt{nh} \left\{ \hat{\beta}(u) - \beta(u) - \frac{h^2}{2} \mu_2(K) \beta''(u) + o_p(h^2) \right\} \xrightarrow{d} \\ N \left(0, \frac{v_0(K)}{f(u)} \Gamma_\delta(u)^{-1} \Omega_\delta(u) \Gamma_\delta(u)^{-1} \right)$$

Its prove process similar to *Wei(2011)* proof of theorem 2.1.

Proof of Theorems 2. Because in the field of $u, \beta_j(\cdot)$ is smooth, by Taylor expansion, we can obtain:

$$X_i^T \beta(U_i) = X_i^T \beta(u) + X_i^T h \beta'(u) \frac{U_i-u}{h} + \frac{h^2}{2} X_i^T \beta''(u) \left(\frac{U_i-u}{h} \right)^2 + o_p(h^2)$$

$$\left(\frac{U_i-u}{h} \right)^2 + o_p(h^2)$$

By the definition of $\tilde{\beta}(u)$, we have

$$\tilde{\beta}(u) - \beta(u) \\ = (I_p O_p) \{D_u^T W_u D_u - \Omega_u\}^{-1} D_u^T W_u (Y^* - C_u) - \beta(u) \\ = (I_p O_p) \{D_u^T W_u D_u - \Omega_u\}^{-1} D_u^T W_u A \frac{h^2 \beta''(u)}{2} \\ + (I_p O_p) \{D_u^T W_u D_u - \Omega_u\}^{-1} D_u^T W_u I_n o_p(h^2) \\ + (I_p O_p) \{D_u^T W_u D_u - \Omega_u\}^{-1} D_u^T W_u (L-\delta) \xi^T \hat{\beta}(u) \\ - (I_p O_p) \{D_u^T W_u D_u - \Omega_u\}^{-1} D_u^T W_u C_u \\ + (I_p O_p) \{D_u^T W_u D_u - \Omega_u\}^{-1} D_u^T W_u e$$

$$+ (I_p O_p) \{D_u^T W_u D_u - \Omega_u\}^{-1} \left(\Omega_u - D_u^T W_u B \right) \theta(u) \\ = I_1 + I_2 + I_3 - I_4 + I_5 + I_6$$

Where, $L = [1, \dots, 1]^T, \delta = (\delta_1, \dots, \delta_n)^T$,

By lemma 1 and lemma 2, it is easy to prove

$$I_1 = \frac{h^2}{2} \mu_2(K) \beta''(u) + o_p(1), I_2 = o_p(h^2), I_3 = I_4$$

$$\text{and } \sqrt{nh} \left\{ \tilde{\beta}(u) - \beta(u) - \frac{h^2}{2} \mu_2(K) \beta''(u) + o_p(h^2) \right\} \\ = f(u)^{-1} \Gamma(u)^{-1} \sqrt{nh} \frac{1}{n} \sum_{i=1}^n K_h(U_i-u) \{V_i \varepsilon_i + \eta_i \beta(u)\} + o_p(1)$$

Therefore, by the central limit theorem can be

$$\sqrt{nh} \left\{ \tilde{\beta}(u) - \beta(u) - \frac{h^2}{2} \mu_2(K) \beta''(u) + o_p(h^2) \right\} \xrightarrow{d} \\ N \left(0, \frac{v_0(K)}{f(u)} \Gamma(u)^{-1} \Omega(u) \Gamma(u)^{-1} \right)$$

REFERENCES

- [1] Hastie, T., Tibshirani, R. Varying-coefficient models. Royal Statistical Society, 1993,55(4):757-796.
- [2] Fan, J., Zhang J. Statistica estimation in varying coefficient models. Ann Statist, 1999,27:1491-1518.
- [3] Chiang, C., Rice, J., Wu, C. Smoothing spline estimation for varying coefficient models with repeatedly measured dependent variables. J Am Stat Assoc, 2001,96:605-619.
- [4] Huang, J. Z., Wu, C. O., Zhou L. Varying-coefficient models and basis function approximations for the analysis of repeated measurements. Biometrika, 2002,89:111-128.
- [5] Hoover, D. R., Rice, J. A., Wu, C. O., Yang, L. P. Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. Biometrika, 1998,85:809-822.
- [6] Wu, C. O., Chiang, C., Hoover, D. R. Asymptotic confidence regions for kernel smoothing of a varying-coefficient model with longitudinal data. J Am Stat Assoc, 1998,93:1388-1402.
- [7] Wei, C. Estimation in partially linear Errors-in-Variables Models with Missing Response Variables. Journal of Mathematical Physics(China), 2010,30A(4):1042-1054.
- [8] Wei, C. Estimation in Varying-Coefficient Errors-in-Variables Models with Missing response Variables. Taylor&Francis Group, 2011,40:383-393.
- [9] Chu, C., Cheng, P. Nonparametric regression estimation with missing data. J Statist Planning Inference, 1995,48:85-99.
- [10] Mack, Y. P., Silverman, B. W. Weak and strong uniform consistency of kernel regression estimates. Z Wahrsch Verw Gebiete, 1982,61:405-415.
- [11] Cai, Z., Fan, J., Yao, Q. Functional coefficient regression models for nonlinear times series. Journal of the American Statistical Association, 2000,95:941-956