Chin. Ann. Math.
32B(5), 2011, 729–740
DOI: 10.1007/s11401-011-0669-9

Chinese Annals of Mathematics, Series B

Efficient Quantile Estimation for Functional-Coefficient Partially Linear Regression Models*

Zhangong ZHOU[1] Rong JIANG[2] Weimin QIAN[3]

Abstract The quantile estimation methods are proposed for functional-coefficient partially linear regression (FCPLR) model by combining nonparametric and functional-coefficient regression (FCR) model. The local linear scheme and the integrated method are used to obtain local quantile estimators of all unknown functions in the FCPLR model. These resulting estimators are asymptotically normal, but each of them has big variance. To reduce variances of these quantile estimators, the one-step backfitting technique is used to obtain the efficient quantile estimators of all unknown functions, and their asymptotic normalities are derived. Two simulated examples are carried out to illustrate the proposed estimation methodology.

Keywords Functional-coefficient model, Quantile regression, Local linear method, Backfitting technique, Asymptotic normality
2000 MR Subject Classification 62J05, 62G08, 62E20

1 Introduction

Consider the following functional-coefficient partially linear regression (FCPLR) model:

$$Y_i = a_0(X_i) + Z_i^{\mathrm{T}} a(U_i) + \varepsilon_i, \tag{1.1}$$

where $\{Y_i, X_i, U_i, Z_i\}_{i=1}^n$ are observations, X_i and U_i are scalar, $Z_i = (Z_{i1}, \cdots, Z_{iq})^{\mathrm{T}}$ is a $q \times 1$ vector, and $\{\varepsilon_i\}_{i=1}^n$ denote the random errors. $a(\cdot)$ is a q-variate vector of unknown smooth functions, $a_0(\cdot)$ is called the constant part function, and $a(\cdot)$ the coefficient function.

Model (1.1) is flexible enough to include many well-studied important parametric, semiparametric and nonparametric regression models. For example, when $a(\cdot)$ is a $q \times 1$ unknown vector, model (1.1) becomes partially linear regression model, which was first introduced by Engle et al. [3], and was systematically studied in [5]. When $a_0 = 0$, model (1.1) is reduced to the varying-coefficient regression model proposed by Hastie and Tibshirani [6], and further was investigated by Cai et al. [1] and Zhang et al. [14] among others. When $a(\cdot)$ is a q-variate vector of unknown functions, model (1.1) becomes functional-coefficient partially linear regression (FCPLR) model proposed by Wong et al. [13]. By using local linear methods and the one-step backfitting technique, they obtained the improved estimators (called the efficient estimator) of $a_0(\cdot)$ and $a(\cdot)$, and it is shown that the resulting estimators are asymptotically normal.

Manuscript received September 20, 2010. Revised June 8, 2011.
[1]Department of Statistics, Jiaxing University, Jiaxing 314001, Zhejiang, China.
E-mail: Zhouzg@mail.zjxu.edu.cn
[2]Department of Mathematics, Tongji University, Shanghai 200092, China. E-mail: jrtrying@126.com
[3]Corresponding author. Department of Mathematics, Tongji University, Shanghai 200092, China.
E-mail: wmqian2003@yahoo.com.cn
*Project supported by the Zhejiang Provincial Natural Science Foundation of China (No. Y6110662).

The existing estimators of the unknown functions in model (1.1) are obtained by the least squares method. The least squares estimates certainly have some nice properties, in particular when the random error follows the normal distribution. As is well-known in the robustness literature, estimation and inference based on least squares are highly sensitive to outliers in the data. Hence, more robust estimation methods are required. To this end, the quantile estimation technique was widely implemented with nonparametric methods to overcome the limitation of the least squares. For example, a varying-coefficient model for the conditional quantile regression was considered recently by several authors. Honda [7] and Kim [8] studied varying-coefficient models for independent data by using local polynomials and splines, respectively. Cai and Xu [2] considered local polynomial estimators for time series data. Mu and Wei [11] studied a dynamic quantile regression transformation model for longitudinal data. Wang et al. [12] developed theory and methodology for analyzing longitudinal data in the quantile partially linear varying coefficient models by using splines. This motivates us to apply the quantile regression method to FCPLR model (1.1).

In this paper, we propose the quantile estimators of the unknown functions in FCPLR model (1.1), and assume that $P(\varepsilon_i \le 0 \mid X_i, U_i, Z_i) = \tau$ for $0 < \tau < 1$. The local linear scheme and the integrated method are used to obtain local quantile estimators of all functions in model (1.1). Under some regularity conditions, these resulting estimators are asymptotically normal, but have big variance. With the aim to reduce variance of the local quantile estimators, the one-step backfitting technique proposed by Linton [10] is adopted to obtain the efficient quantile estimators of all functions. Furthermore, it is shown that the efficient quantile estimator of the constant part function has the same asymptotic normality property as the local linear estimator for univariate nonparametric quantile regression model and these efficient quantile estimators of coefficient functions share the same asymptotic normality properties as local linear estimators for the varying-coefficient quantile regression model (see [2]).

The rest of this paper is organized as follows. The quantile estimation methods for the constant part function and the coefficient functions are proposed in Section 2, and some assumptions and the asymptotic properties of the proposed estimators are presented in this section. In Section 3, some simulations are conducted to demonstrate the proposed estimation methodology. Proofs of the theorems are presented in Section 4.

2 Estimation

Suppose that $\{Y_i, X_i, U_i, Z_i\}_{i=1}^n$ is a random sample from model (1.1). Throughout the article, $K_\alpha(\cdot)$ is a bounded, compactly supported symmetric about zero and Lipschitz continuous density function, and $h_\alpha > 0$ is a bandwidth, $\alpha = 1, 2, 3, 4$. We also assume that $a_0(\cdot)$ and $a(\cdot)$ have Lipschitz continuous second derivatives.

2.1 Local quantile estimator

We estimate $a_0(\cdot)$ and $a(\cdot)$ by using the local linear method (see [4]) based on observations $\{Y_i, X_i, U_i, Z_i\}_{i=1}^n$. Approximate $a_0(\cdot)$ and $a(\cdot)$ in the neighbors of x_0 and u_0 by $a_0(x_0) + a_0'(x_0)(x - x_0)$ and $a(u_0) + a'(u_0)(u - u_0)$, respectively. We minimize the weighted loss function

$$\sum_{i=1}^n \rho_\tau(Y_i - \alpha_0 - Z_i^{\mathrm{T}}\beta_0 - \alpha_1(X_i - x_0) - Z_i^{\mathrm{T}}\beta_1(U_i - u_0))K_1\Big(\frac{X_i - x_0}{h_1}\Big)K_2\Big(\frac{U_i - u_0}{h_2}\Big) \tag{2.1}$$

with respect to α_j and β_j for $j = 0, 1$, where $\rho_\tau(u) = u(\tau - I(u < 0))$ is the quantile loss function for $0 < \tau < 1$. Solving the minimization problem in (2.1) gives the initial estimator $(\widehat{\alpha}_0(x_0, u_0), \widehat{\alpha}_1(x_0, u_0), \widehat{\beta}_0^{\mathrm{T}}(x_0, u_0), \widehat{\beta}_1^{\mathrm{T}}(x_0, u_0))^{\mathrm{T}}$ for $(a_0(x_0), a_0'(x_0), a^{\mathrm{T}}(u_0), a'^{\mathrm{T}}(u_0))^{\mathrm{T}}$.

By the integrated method, local quantile estimator of the constant part function $a_0(x_0)$ is defined as

$$\widehat{a}_0(x_0) = \frac{1}{n}\sum_{j=1}^{n} \widehat{\alpha}_0(x_0, U_j). \tag{2.2}$$

Local quantile estimator of the coefficient part function $a(u_0)$ is defined as

$$\widehat{a}(u_0) = \frac{1}{n}\sum_{k=1}^{n} \widehat{\beta}_0(X_k, u_0). \tag{2.3}$$

In what follows, define

$$\begin{aligned}
v_l &= E(f_{\varepsilon\,|\,X_1,U_1,Z_1}(0) Z_1^l \mid X_1 = x_0, U_1 = u_0) \quad \text{for } l = 0, 1,\\
v_2 &= E(f_{\varepsilon\,|\,X_1,U_1,Z_1}(0) Z_1 Z_1^{\mathrm{T}} \mid X_1 = x_0, U_1 = u_0),\\
w_1 &= E(Z_1 \mid X_1 = x_0, U_1 = u_0),\\
w_2 &= E(Z_1 Z_1^{\mathrm{T}} \mid X_1 = x_0, U_1 = u_0),\\
\mu_2^{(j)} &= \int u^2 K_j(u)\mathrm{d}u, \quad \nu_l^{(j)} = \int u^l K_j^2(u)\mathrm{d}u,
\end{aligned}$$

where $f_{\varepsilon\,|\,x,u,z}(\,\cdot\,)$ is the conditional density of ε given X, U and Z. Also, set

$$\Sigma_1(x_0, u_0) = \begin{pmatrix} v_0 & v_1^{\mathrm{T}} & 0 & 0\\ v_1 & v_2 & 0 & 0\\ 0 & 0 & \mu_2^{(1)} v_0 & 0\\ 0 & 0 & 0 & \mu_2^{(2)} v_2 \end{pmatrix},$$

$$\Sigma_2(x_0, u_0) = \begin{pmatrix} \nu_0^{(1)}\nu_0^{(2)} & \nu_0^{(1)}\nu_0^{(2)} w_1^{\mathrm{T}} & \nu_1^{(1)}\nu_0^{(2)} & \nu_0^{(1)}\nu_1^{(2)} w_1^{\mathrm{T}}\\ \nu_0^{(1)}\nu_0^{(2)} w_1 & \nu_1^{(1)}\nu_0^{(2)} w_2 & \nu_1^{(1)}\nu_1^{(2)} w_1 & \nu_0^{(1)}\nu_1^{(2)} w_2^{\mathrm{T}}\\ \nu_1^{(1)}\nu_0^{(2)} & \nu_1^{(1)}\nu_1^{(2)} w_1^{\mathrm{T}} & \nu_2^{(1)}\nu_0^{(2)} & \nu_1^{(1)}\nu_1^{(2)} w_1^{\mathrm{T}}\\ \nu_0^{(1)}\nu_1^{(2)} w_1 & \nu_0^{(1)}\nu_1^{(2)} w_2 & \nu_1^{(1)}\nu_1^{(2)} w_1 & \nu_0^{(1)}\nu_2^{(2)} w_2 \end{pmatrix},$$

$$\begin{pmatrix} v_0 & v_1^{\mathrm{T}}\\ v_1 & v_2 \end{pmatrix}^{-1} = \begin{pmatrix} \gamma_{11}(x_0, u_0) & \gamma_{12}(x_0, u_0)\\ \gamma_{21}(x_0, u_0) & \gamma_{22}(x_0, u_0) \end{pmatrix} \quad \text{and} \quad \Omega(x_0, u_0) = \begin{pmatrix} 1 & w_1^{\mathrm{T}}\\ w_1 & w_2 \end{pmatrix}.$$

Let $f_x(x)$, $f_u(u)$ and $f_{xu}(x, u)$ be the density of X, U and (X, U), respectively. Throughout this section, we use $c > 0$ generically to represent any constant which may take a different value for each appearance. To obtain our results, the following assumptions are needed:

(C1) $f_{\varepsilon\,|\,x,u,z}(\,\cdot\,)$ is bounded, Lipschitz continuous and $f_{\varepsilon\,|\,x,u,z}(0) \ge c > 0$.

(C2) $\Sigma_1(x_0, u_0)$ is positive-definite and continuous in the neighborhood of (x_0, u_0).

(C3) $\Omega(x_0, u_0)$ is positive-definite and continuous in the neighborhood of (x_0, u_0).

(C4) There is an $s > 2$ such that $E\|Z_1\|^{2s} < \infty$.

Write $\psi_\tau(x) = \tau - \mathrm{I}_{\{x<0\}}$, $X_{ih_1} = \frac{X_i - x_0}{h_1}$, $U_{ih_2} = \frac{U_i - u_0}{h_2}$, $V_i = (1, Z_i^{\mathrm{T}}, X_{ih_1}, U_{ih_2} Z_i^{\mathrm{T}})^{\mathrm{T}}$, $r_i \equiv r(X_i, U_i) = a_0(X_i) - a_0(x_0) - a_0'(x_0)(X_i - x_0) + Z_i^{\mathrm{T}}[a(U_i) - a(u_0) - a'(u_0)(U_i - u_0)]$, $\Theta = \sqrt{nh_1h_2}(\alpha_0 - a_0(x_0), \beta_0^{\mathrm{T}} - a^{\mathrm{T}}(u_0), h_1(\alpha_1 - a_0'(x_0)), h_2(\beta_1 - a'(u_0))^{\mathrm{T}})^{\mathrm{T}}$. Denote an estimator for Θ by $\widehat{\Theta}$.

The following result states the asymptotic normality of the initial estimator.

Theorem 2.1 *Under assumptions* (C1)–(C4), *if* $nh_1h_2 \to \infty$, $h_1 \to 0$ *and* $h_2 \to 0$, *we have*

$$\sqrt{nh_1h_2}\left[\begin{pmatrix} \widehat{\alpha}_0 - a_0(x_0) \\ \widehat{\beta}_0 - a(u_0) \\ h_1(\widehat{\alpha}_1 - a_0'(x_0)) \\ h_2(\widehat{\beta}_1 - a'(u_0)) \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \mu_2^{(1)} h_1^2 a_0''(x_0) \\ \mu_2^{(2)} h_2^2 a''(u_0) \\ 0 \\ 0 \end{pmatrix}(1 + o_p(1))\right] \xrightarrow{\mathrm{d}} N(0, \Sigma),$$

where $\Sigma = \tau(1-\tau)\Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}\frac{1}{f(x_0,u_0)}$ *and* $\xrightarrow{\mathrm{d}}$ *stands for convergence in distribution.*

The following two theorems give the asymptotic normalities of the local quantile estimators.

Theorem 2.2 *Under assumptions* (C1)–(C4), *and with* $nh_1h_2 \to \infty$, $h_1 = c_1 n^{-\frac{1}{5}}$, $nh_2^5 \to 0$, *where* c_1 *is a positive constant, we have*

$$\sqrt{nh_1}\Big(\widehat{a}_0(x_0) - a_0(x_0) - \frac{1}{2}h_1^2\mu_2^{(1)}a_0''(x_0)\Big) \xrightarrow{\mathrm{d}} N(0, \tau(1-\tau)\nu_0^{(1)}\omega_0(x_0)), \tag{2.4}$$

where $\omega_0(x_0) = \int \frac{f_u^2(u)}{f_{xu}(x_0,u)}\gamma_1^{\mathrm{T}}(x_0,u)\Omega(x_0,u)\gamma_1(x_0,u)\mathrm{d}u$ *and* $\gamma_1^{\mathrm{T}}(x_0,u) = (\gamma_{11}(x_0,u), \gamma_{12}(x_0,u))$.

Theorem 2.3 *Under assumptions* (C1)–(C4), *and with* $nh_1h_2 \to \infty$, $h_2 = c_2 n^{-\frac{1}{5}}$, $nh_1^5 \to 0$, *where* c_2 *is a positive constant, we have*

$$\sqrt{nh_2}\Big(\widehat{a}(u_0) - a(u_0) - \frac{1}{2}h_2^2\mu_2^{(2)}a''(u_0)\Big) \xrightarrow{\mathrm{d}} N(0, \tau(1-\tau)\nu_0^{(2)}\omega(u_0)), \tag{2.5}$$

where $\omega(u_0) = \int \frac{f_x^2(x)}{f_{xu}(x,u_0)}\gamma_2^{\mathrm{T}}(x,u_0)\Omega(x,u_0)\gamma_2(x,u_0)\mathrm{d}x$ *and* $\gamma_2^{\mathrm{T}}(x,u_0) = (\gamma_{21}(x,u_0), \gamma_{22}(x,u_0))$.

Remark 2.1 To discuss how the local quantile estimator $\widehat{a}(u_0)$ works, we consider the functional-coefficient regression model as $Y_i = Z_i^{\mathrm{T}}a(U_i) + \varepsilon_i$, where $P(\varepsilon_i \le 0 \mid X_i, U_i, Z_i) = \tau$ for $0 < \tau < 1$. Under fairly standard assumptions used in [2], the local quantile estimator $\overline{a}(u_0)$ for $a(\cdot)$ satisfies

$$\sqrt{nh_2}\Big(\overline{a}(u_0) - a(u_0) - \frac{h_2^2}{2}\mu_2^{(2)}a''(u_0)\Big) \xrightarrow{\mathrm{d}} N(0, \tau(1-\tau)\nu_0^{(2)}\Sigma_a(u_0)), \tag{2.6}$$

where $\Sigma_a(u_0) = [\Sigma_0^*(u_0)]^{-1}\Sigma_0(u_0)[\Sigma_0^*(u_0)]^{-1}\frac{1}{f_u(u_0)}$ with $\Sigma_0(u_0) = E[Z_1Z_1^{\mathrm{T}} \mid U_1 = u_0]$ and $\Sigma_0^*(u_0) = E[f_{\varepsilon|U_1,Z_1}(0)Z_1Z_1^{\mathrm{T}} \mid U_1 = u_0]$.

From (2.5) and (2.6), we can see that $\widehat{a}(u_0)$ has the same bias as $\overline{a}(u_0)$. To compare their variance, for simplicity, we suppose that $E(Z) = 0$, and X, Z and U are independent of each other, which implies $\gamma_{22}(x_0,u_0) = \Sigma_0^{*-1}(u_0)$ and $w_2 = \Sigma_0(u_0)$ by simple calculation. Hence,

$$\omega(u_0) = \gamma_2^{\mathrm{T}}(x,u_0)\Omega(x,u_0)\gamma_2(x,u_0)\frac{1}{f_u(u_0)} \ge \gamma_{22}(x_0,u_0)w_2\gamma_{22}(x_0,u_0)\frac{1}{f_u(u_0)} = \Sigma_a(u_0),$$

so that the variance of $\widehat{a}(u_0)$ is bigger than that of $\overline{a}(u_0)$. Similarly, we can discuss the local quantile estimator $\widehat{a}_0(x_0)$.

2.2 Efficient quantile estimator

To deal with the asymptotic variances for the proposed local quantile estimators, we adopt the local linear smoothing and one-step backfitting technique to obtain the efficient quantile estimators for all unknown functions.

The efficient quantile estimator for $a(u_0)$ is defined as $\widehat{a}^{\mathrm{e}}(u_0) = (\mathrm{I}_q\ 0_q)\widehat{\beta}^{\mathrm{e}}$, where the vector $\widehat{\beta}^{\mathrm{e}} = (\widehat{b}_0^{\mathrm{T}}, \widehat{b}_1^{\mathrm{T}})^{\mathrm{T}}$ minimizes

$$\sum_{i=1}^{n}\rho_\tau\left(Y_i - Z_i^{\mathrm{T}}[b_0 + b_1(U_i - u_0)] - \widehat{a}_0(X_i)\right)K_3\Big(\frac{U_i - u_0}{h_3}\Big) \tag{2.7}$$

with respect to b_0 and b_1.

As to $a_0(\cdot)$, the efficient quantile estimator can be defined as $\widehat{a}_0^{\rm e}(x_0) = (1,0)\widehat{\theta}^{\rm e}$, where the vector $\widehat{\theta}^{\rm e} = (\widehat{\theta}_0, \widehat{\theta}_1)^{\rm T}$ minimizes

$$\sum_{i=1}^{n} \rho_\tau \left(Y_i - \theta_0 - \theta_1(X_i - x_0) - Z_i^{\rm T}\widehat{a}(U_i)\right) K_4\Big(\frac{X_i - x_0}{h_4}\Big) \tag{2.8}$$

with respect to θ_0 and θ_1.

To derive our results, the following assumptions are also needed:

(C5) The conditional density $f_{\varepsilon\,|\,u,z}(\cdot)$ of ε given U and Z is bounded, Lipschitz continuous and $f_{\varepsilon\,|\,u,z}(0) \geq c > 0$.

(C6) $\Sigma_0(u_0) = E[Z_1 Z_1^{\rm T} \mid U_1 = u_0]$ and $\Sigma_0^*(u_0) = E[f_{\varepsilon\,|\,U_1,Z_1}(0) Z_1 Z_1^{\rm T} \mid U_1 = u_0]$ are positive definite and continuous in neighborhood of u_0.

(C7) The conditional density $f_{\varepsilon|x}(\cdot)$ of ε given X is bounded, Lipschitz continuous and $f_{\varepsilon\,|\,x}(0) \geq c > 0$.

The following results state the asymptotic normalities of the efficient quantile estimators.

Theorem 2.4 *Under assumptions* (C1)–(C6)*, if* $nh_1h_2 \to \infty$*,* $\frac{h_1}{h_3} \to 0$*,* $\frac{h_2}{h_3} \to 0$ *and* $h_3 = c_3 n^{-\frac{1}{5}}$*, where* c_3 *is a positive constant, we have*

$$\sqrt{nh_3}\Big(\widehat{a}^{\rm e}(u_0) - a(u_0) - \frac{h_3^2}{2}\mu_2^{(3)} a''(u_0)\Big) \xrightarrow{\rm d} N(0, \tau(1-\tau)\nu_0^{(3)}\Sigma_a(u_0)), \tag{2.9}$$

where $\Sigma_a(u_0) = [\Sigma_0^*(u_0)]^{-1}\Sigma_0(u_0)[\Sigma_0^*(u_0)]^{-1}\frac{1}{f_u(u_0)}$.

Theorem 2.5 *Under assumptions* (C1)–(C4) *and* (C7)*, if* $nh_1h_2 \to \infty$*,* $\frac{h_1}{h_4} \to 0$*,* $\frac{h_2}{h_4} \to 0$ *and* $h_4 = c_4 n^{-\frac{1}{5}}$*, where* c_4 *is a positive constant, we have*

$$\sqrt{nh_4}\Big(\widehat{a}_0^{\rm e}(x_0) - a_0(x_0) - \frac{h_4^2}{2}\mu_2^{(4)} a_0''(x_0)\Big) \xrightarrow{\rm d} N(0, \sigma_0^2(x_0)), \tag{2.10}$$

where $\sigma_0^2(x_0) = \frac{\tau(1-\tau)\nu_0^{(4)}}{f_{\varepsilon|x_0}^2(0) f_x(x_0)}$.

Remark 2.2 From Theorems 2.4–2.5, we can see that the efficient quantile estimator of the constant part function has the same asymptotic normality property as the local linear estimator for univariate nonparametric quantile regression model and these efficient quantile estimators of coefficient functions share the same asymptotic normality properties as local linear estimators for the varying-coefficient quantile regression model (see [2]). Moreover, our proposed estimation method will also apply if the data are strictly α-mixing and stationary.

3 Simulations

To illustrate the advantage of the quantile estimators in this section, we present Monte Carlo simulations to compare the proposed estimation methods with the least squares (LS) method in [13].

The efficient quantile estimators $\{\widehat{a}_j(\cdot)\}_{j=0}^{q}$ are computed via the mean absolute deviation errors (MADE), defined as $\text{MADE}_j = n_j^{-1}\sum\limits_{k=1}^{n_j} |\widehat{a}_j(v_k) - a_j(v_k)|$, and $\{v_k = 0.05k - 1 : 1 \leq k \leq n_j = 38\}$ are the grid points. We use the Epanechnikov kernel $K(u) = 0.75(1 - u^2 {\rm I}_{(|u|<1)})$ for every $K_\alpha(\cdot)$, $\alpha = 1, 2, 3, 4$. For each replication, sample size and τ, bandwidths h_1 and h_2 can be selected by minimizing the following cross-validation score:

$$CV(h_1, h_2) = \sum_{i=1}^{n} \rho_\tau (Y_i - \widehat{a}_0^{(-i)}(X_i) - Z_i^{\mathrm{T}} \widehat{a}^{(-i)}(U_i)),$$

where $\widehat{a}_0^{(-i)}$ and $\widehat{a}^{(-i)}$ are local quantile estimates after deleting the i-th subject. Similarly, bandwidths h_3 and h_4 are also selected by this method.

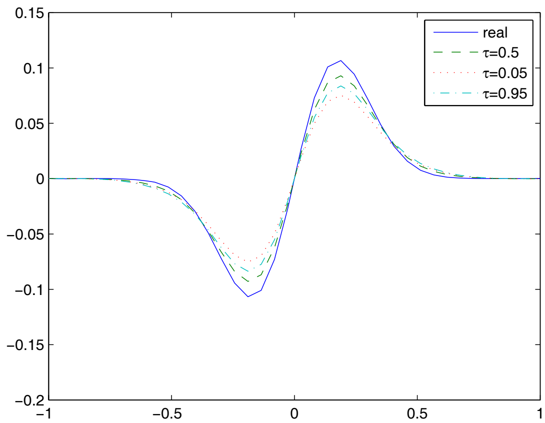
Example 3.1 The data are generated from the functional-coefficient partially linear regression model $Y = a_0(X) + a_1(U)Z_1 + a_2(U)Z_2 + (\varepsilon - F^{-1}(\tau))$, where $a_0(x) = x\exp(-16x^2)$, $a_1(u) = u\exp(-3u^2)$, $a_2(u) = u^2$, $(X, U, Z_1, Z_2)^{\mathrm{T}} \sim U([-1, 1]^4)$, $\varepsilon \sim t(3)$ and $F^{-1}(\tau)$ is the τ-th quantile of $t(3)$, where $t(3)$ is a t-distribution with 3 degrees of freedom.

Example 3.2 The data are generated from the functional-coefficient partially linear regression model $Y_t = a_0(X_t) + a_1(Z_{t-1})Z_t + a_2(Z_{t-1})Z_{t-2} + (\epsilon_t - F_\epsilon^{-1}(\tau))$, where $a_0(x), a_1(u)$ and $a_2(u)$ are the same as those in Example 3.1; $X_t = 1.1X_{t-1} - 0.3X_{t-2} + e_t$, e_t are i.i.d., $N(0, 0.2^2)$, $Z_t = -0.6Z_{t-1} - 0.3Z_{t-2} + \zeta_t$, ζ_t are i.i.d., $N(0, 0.3^2)$, $\epsilon_t \sim U(-0.5, 0.5)$ and $F_\epsilon^{-1}(\tau)$ is the τ-th quantile of $U(-0.5, 0.5)$.

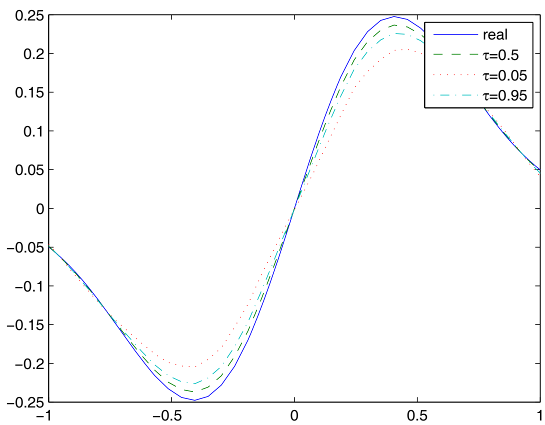
The Monte Carlo simulation is repeated 500 times for each sample size $n = 200, 300$ and 400, and for each $\tau = 0.05, 0.5$ and 0.95. We compute the median and standard deviation of 500 MADE values for each scenario and summarize the results in Table 1. In addition, Figure 1 presents the quantile estimates of $a_0(\cdot)$, $a_1(\cdot)$ and $a_2(\cdot)$, which are drawn from Example 3.1 with $\tau = 0.05, 0.5, 0.95$ and a sample size 200.

Table 1 The median and standard deviation of 500 MADE values

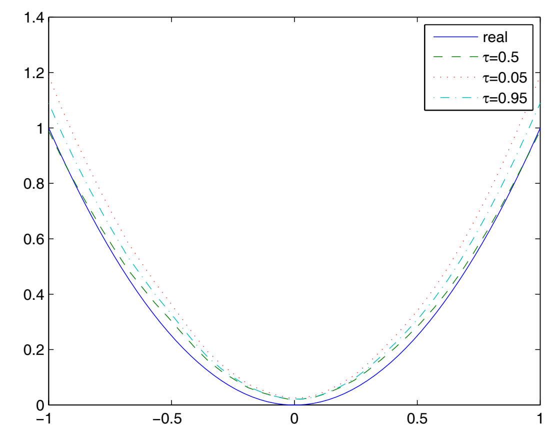
		Example 1		
	n	200	300	400
LS	MADE_0	0.1156 (0.0927)	0.0978 (0.0632)	0.0706 (0.0526)
	MADE_1	0.1073 (0.0325)	0.0936 (0.0338)	0.0684 (0.0342)
	MADE_2	0.1042 (0.0382)	0.0955 (0.0366)	0.0691 (0.0388)
$\tau = 0.05$	MADE_0	0.1082 (0.0912)	0.0922 (0.0621)	0.0646 (0.0508)
	MADE_1	0.1022 (0.0320)	0.0903 (0.0310)	0.0621 (0.0312)
	MADE_2	0.1002 (0.0342)	0.0925 (0.0328)	0.0646 (0.0356)
$\tau = 0.5$	MADE_0	0.0592 (0.0216)	0.0461 (0.0205)	0.0432 (0.0188)
	MADE_1	0.0518 (0.0261)	0.0398 (0.0256)	0.0356 (0.0246)
	MADE_2	0.0520 (0.0272)	0.0384 (0.0266)	0.0371 (0.0252)
$\tau = 0.95$	MADE_0	0.1068 (0.0910)	0.0918 (0.0614)	0.0637 (0.0513)
	MADE_1	0.1013 (0.0318)	0.0898 (0.0308)	0.0619 (0.0314)
	MADE_2	0.0098 (0.0339)	0.0916 (0.0314)	0.0639 (0.0349)
		Example 2		
LS	MADE_0	0.1190 (0.0936)	0.1066 (0.0622)	0.0828 (0.0530)
	MADE_1	0.1108 (0.0538)	0.1102 (0.0425)	0.0880 (0.0432)
	MADE_2	0.1072 (0.0568)	0.1090 (0.0416)	0.0836 (0.0428)
$\tau = 0.05$	MADE_0	0.1102 (0.0920)	0.1002 (0.0634)	0.0716 (0.0528)
	MADE_1	0.1056 (0.0562)	0.1041 (0.0412)	0.0813 (0.0425)
	MADE_2	0.1038 (0.0480)	0.0983 (0.0407)	0.0722 (0.0402)
$\tau = 0.5$	MADE_0	0.0752 (0.0430)	0.0602 (0.0346)	0.0596 (0.0292)
	MADE_1	0.0718 (0.0469)	0.0669 (0.0351)	0.0558 (0.0360)
	MADE_2	0.0716 (0.0420)	0.0586 (0.0378)	0.0550 (0.0298)
$\tau = 0.95$	MADE_0	0.1093 (0.0860)	0.0938 (0.0623)	0.0710 (0.0462)
	MADE_1	0.1049 (0.0540)	0.0913 (0.0415)	0.0812 (0.0409)
	MADE_2	0.1012 (0.0500)	0.0972 (0.0406)	0.0718 (0.0398)



(a)



(b)



(c)

Figure 1 (a), (b) and (c) present the quantile estimates of $a_0(\cdot)$, $a_1(\cdot)$ and $a_2(\cdot)$, respectively.

From Table 1, we can observe that under non-normal model errors and the different data, the MADE values for the efficient quantile estimates decrease as n increases for all three values of τ, and our estimators are better than the ones based on the least squares method, which imply that the proposed method is very efficient. Also, the performance for the median quantile estimate is slightly better than that for two tails ($\tau = 0.05$ and 0.95).

Figure 1 also implies that the proposed estimators $\widehat{a}_0(\cdot)$, $\widehat{a}_1(\cdot)$ and $\widehat{a}_2(\cdot)$ work well, because the four lines are very close to each other. Overall speaking, the proposed procedure is reliable and performs fairly well.

4 Proofs of the Main Results

To prove the main results, we need the following two lemmas whose proofs can be found in [9].

Lemma 4.1 *Let $V_n(\theta)$ be a vector function that satisfies*

(1) $-\theta^{\mathrm{T}}V_n(\lambda\theta) \geq -\theta^{\mathrm{T}}V_n(\theta)$ *for* $\lambda \geq 1$;

(2) $\sup\limits_{\|\theta\|\leq M} \|V_n(\theta) + D\theta - A_n\| = o_p(1)$,

where $\|A_n\| = O_p(1)$, $0 < M < \infty$, *and D is a positive-definite matrix. Suppose that θ_n is a vector such that* $\|V_n(\theta_n)\| = o_p(1)$. *Then* $\|\theta_n\| = O_p(1)$ *and* $\theta_n = D^{-1}A_n + o_p(1)$.

Lemma 4.2 *Let $\widehat{\beta}$ be the minimizer of the function*

$$\sum_{i=1}^{n} \omega_i\rho_\tau(Y_i - V_i^{\mathrm{T}}\beta),$$

where $\omega_i > 0$. *Then*

$$\sum_{i=1}^{n} \omega_i\psi_\tau(Y_i - V_i^{\mathrm{T}}\widehat{\beta}) \leq \dim(V_1) \max_{1\leq i\leq n} \|\omega_i V_i\|.$$

By the definition of Θ in Section 2.1, we have

$$\begin{pmatrix}\alpha_0\\ \beta_0\\ \alpha_1\\ \beta_1\end{pmatrix} = \begin{pmatrix}a_0(x_0)\\ a(u_0)\\ a_0'(x_0)\\ a'(u_0)\end{pmatrix} + \delta_n\Theta,$$

where $\delta_n = (nh_1h_2)^{-\frac{1}{2}}$. Thus, $Y_i - \alpha_0 - \alpha_1(X_i - x_0) - Z_i^{\mathrm{T}}\{\beta_0 + \beta_1(U_i - u_0)\} = \varepsilon_i + r_i - \delta_n\Theta^{\mathrm{T}}V_i$. Therefore,

$$\widehat{\Theta} = \arg\min_{\Theta} \sum_{i=1}^{n} \rho_\tau(\varepsilon_i + r_i - \delta_n\Theta^{\mathrm{T}}V_i)K_1(X_{ih_1})K_2(U_{ih_2}).$$

Now, define $V_n(\Theta) = \delta_n \sum\limits_{i=1}^{n} \psi_\tau(\varepsilon_i + r_i - \delta_n\Theta^{\mathrm{T}}V_i)V_iK_1(X_{ih_1})K_2(U_{ih_2})$ and $\mathbf{B}_M = \{\|\Theta\| \leq M\}$ for any $0 < M < \infty$.

Lemma 4.3 *Under the assumptions of Theorem* 2.1, *we have*

(1) $\sup\limits_{\Theta\in\mathbf{B}_M} \|V_n(\Theta) - V_n(0) - E[V_n(\Theta) - V_n(0)]\| = o_p(1)$;

(2) $\sup\limits_{\Theta\in\mathbf{B}_M} \|E[V_n(\Theta) - V_n(0)] + f_{xu}(x_0, u_0)\Sigma_1(x_0, u_0)\Theta\| = o(1)$;

(3) $\|V_n(0)\| = O_p(1)$,

where $\Sigma_1(x_0, u_0)$ *is defined in Section* 2.1.

Proof By using the arguments similar to [2, Theorem 1], we can prove Lemma 4.3. Here, we omit the process of the proof.

Proof of Theorem 2.1 By using Lemmas 4.1–4.3 and the arguments similar to [2, Theorem 1], we have

$$\begin{aligned}\widehat{\Theta} &= \frac{\Sigma_1^{-1}(x_0,u_0)}{\sqrt{nh_1h_2}f_{xu}(x_0,u_0)}\sum_{i=1}^{n}\psi_\tau(\varepsilon_i+r_i)V_iK_1(X_{ih_1})K_2(U_{ih_2})+o_p(1)\\ &= \frac{\Sigma_1^{-1}(x_0,u_0)}{\sqrt{nh_1h_2}f_{xu}(x_0,u_0)}\sum_{i=1}^{n}[\psi_\tau(\varepsilon_i+r_i)-\psi_\tau(\varepsilon_i)]V_iK_1(X_{ih_1})K_2(U_{ih_2})\\ &\quad+\frac{\Sigma_1^{-1}(x_0,u_0)}{\sqrt{nh_1h_2}f_{xu}(x_0,u_0)}\sum_{i=1}^{n}\psi_\tau(\varepsilon_i)V_iK_1(X_{ih_1})K_2(U_{ih_2})+o_p(1)\\ &= \mathbf{B}_n+\eta_n+o_p(1). \end{aligned}\tag{4.1}$$

Observe that $E\psi_\tau(\varepsilon_i)=0$ and $\mathrm{Var}[\psi_\tau(\varepsilon_i)]=\tau(1-\tau)$. By using Cramér-Wold device and the similar arguments in [1], one can show that

$$\eta_n \xrightarrow{\mathrm{d}} N\Big(0,\frac{\tau(1-\tau)}{f_{xu}(x_0,u_0)}\Sigma_1^{-1}(x_0,u_0)\Sigma_2(x_0,u_0)\Sigma_1^{-1}(x_0,u_0)\Big).\tag{4.2}$$

By Taylor expansion and simple calculation, we have

$$E[\mathbf{B}_n]=\frac{\sqrt{nh_1h_2}}{2}\begin{pmatrix}\mu_2^{(1)}h_1^2a_0''(x_0)\\ \mu_2^{(2)}h_2^2a''(u_0)\\ 0\\ 0\end{pmatrix}(1+o(1)).\tag{4.3}$$

Since $\psi_\tau(\varepsilon_i+r_i)-\psi_\tau(\varepsilon_i)=\mathrm{I}_{\{\varepsilon_i<0\}}-\mathrm{I}_{\{\varepsilon_i<-r(X_i,U_i)\}}$, we have

$$[\psi_\tau(\varepsilon_i+r_i)-\psi_\tau(\varepsilon_i)]^2=\mathrm{I}_{\{d_{1i}<\varepsilon_i\le d_{2i}\}},$$

where $d_{1i}=\min(0,-r_i)$ and $d_{2i}=\max(0,-r_i)$. Furthermore, simple calculation yields that

$$E[\{\psi_\tau(\varepsilon_i+r_i)-\psi_\tau(\varepsilon_i)\}^2V_iV_i^{\mathrm{T}}K_1^2(X_{ih_1})K_2^2(U_{ih_2})]=O(h_1^3h_2+h_1h_2^3).$$

Therefore, $\mathrm{Var}(\mathbf{B}_n)=o(1)$. This, together with (4.1)–(4.3) and Slutsky Theorem, completes the proof of Theorem 2.1.

Proof of Theorem 2.2 From (4.1), we can obtain

$$\widehat{a}_0(x_0)-a_0(x_0)=(\mathrm{I}_1+\mathrm{I}_2)(1+o_p(1)),\tag{4.4}$$

where

$$\begin{aligned}\mathrm{I}_1&=\frac{1}{n^2h_1h_2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\psi_\tau(\varepsilon_i)}{f_{xu}(x_0,U_j)}\gamma_1^{\mathrm{T}}(x_0,U_j)\begin{pmatrix}1\\ Z_i\end{pmatrix}K_1\Big(\frac{X_i-x_0}{h_1}\Big)K_2\Big(\frac{U_i-U_j}{h_2}\Big),\\ \mathrm{I}_2&=\frac{1}{n^2h_1h_2}\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{\psi_\tau(\varepsilon_i+r_i)-\psi_\tau(\varepsilon_i)}{f_{xu}(x_0,U_j)}\gamma_1^{\mathrm{T}}(x_0,U_j)\begin{pmatrix}1\\ Z_i\end{pmatrix}K_1\Big(\frac{X_i-x_0}{h_1}\Big)K_2\Big(\frac{U_i-U_j}{h_2}\Big).\end{aligned}$$

Similar to the proof of Theorem 1 in [13], I_1 can be represented as

$$\mathrm{I}_1=\mathrm{I}_{11}(1+o_p(1)),\tag{4.5}$$

where

$$\mathrm{I}_{11}=\frac{1}{nh_1}\sum_{i=1}^{n}\frac{f_u(U_i)}{f_{xu}(x_0,U_i)}\psi_\tau(\varepsilon_i)\gamma_1^{\mathrm{T}}(x_0,U_i)\begin{pmatrix}1\\Z_i\end{pmatrix}K_1\Big(\frac{X_i-x_0}{h_1}\Big).$$

Obviously, I_{11} is asymptotically normal with $E(\mathrm{I}_{11})=0$, and variance

$$\begin{aligned}E(\mathrm{I}_{11}^2)&=\frac{\tau(1-\tau)}{nh_1^2}E\Big[\frac{f_u(U_1)}{f_{xu}(x_0,U_1)}\gamma_1^{\mathrm{T}}(x_0,U_1)\begin{pmatrix}1\\Z_1\end{pmatrix}K_1\Big(\frac{X_1-x_0}{h_1}\Big)\Big]^2\\&=\frac{\tau(1-\tau)}{nh_1}\nu_0^{(1)}\omega_0(x_0)(1+o(1)).\end{aligned}$$

Therefore,

$$\sqrt{nh_1}\,\mathrm{I}_{11}\xrightarrow{\mathrm{d}}N(0,\tau(1-\tau)\nu_0^{(1)}\omega_0(x_0)).\tag{4.6}$$

Similar to I_1, I_2 can be written as

$$\mathrm{I}_2=\frac{1}{nh_1}\sum_{i=1}^{n}\frac{\psi_\tau(\varepsilon_i+r_i)-\psi_\tau(\varepsilon_i)}{f_{xu}(x_0,U_i)}f_u(U_i)\gamma_1^{\mathrm{T}}(x_0,U_i)\begin{pmatrix}1\\Z_i\end{pmatrix}K_1\Big(\frac{X_i-x_0}{h_1}\Big)(1+o_p(1)).$$

By the proof of Theorem 2.1, and using that $\gamma_{11}(x_0,u_0)v_0+\gamma_{12}(x_0,u_0)v_1=1$ and $\int_{-\infty}^{\infty}f_u(u)\mathrm{d}u=1$, we can prove that

$$\mathrm{I}_2=\frac{1}{2}h_1^2\mu_2^{(1)}a_0''(x_0)(1+o_p(1)).\tag{4.7}$$

This, together with (4.4)–(4.7), completes the proof of Theorem 2.2.

Proof of Theorem 2.3 By using the similar argument as in Theorem 2.2, we can prove Theorem 2.3. Here, we omit the process of the proof.

Proof of Theorem 2.4 Substituting model (1.1) into the expression (2.7) and using the Taylor expansion, we have

$$\begin{aligned}(\widehat{b}_0^{\mathrm{T}},\widehat{b}_1^{\mathrm{T}})^{\mathrm{T}}=\arg\min_{b_0,b_1}\sum_{i=1}^{n}\rho_\tau\Big\{&\frac{1}{2}Z_i^{\mathrm{T}}a''(\zeta_i)(U_i-u_0)^2+a_0(X_i)-\widehat{a}_0(X_i)+\varepsilon_i\\&-[(b_0-a(u_0))^{\mathrm{T}}+(b_1-a'(u_0))^{\mathrm{T}}(U_i-u_0)]Z_i\Big\}K_3\Big(\frac{U_i-u_0}{h_3}\Big),\end{aligned}$$

where $|\zeta_i-u_0|<|U_i-u_0|$. Write $W_i=\frac{1}{2}Z_i^{\mathrm{T}}a''(\zeta_i)(U_i-u_0)^2$, $\vartheta=(\vartheta_1^{\mathrm{T}},\vartheta_2^{\mathrm{T}})^{\mathrm{T}}=\sqrt{nh_3}\,((b_0-a(u_0))^{\mathrm{T}},h_3(b_1-a'(u_0))^{\mathrm{T}})^{\mathrm{T}}$, $\widetilde{U}_i=(nh_3)^{-\frac{1}{2}}(Z_i^{\mathrm{T}},h_3^{-1}(U_i-u_0)Z_i^{\mathrm{T}})^{\mathrm{T}}$. Then we have the new optimization problem

$$\widehat{\vartheta}^{\mathrm{T}}=\arg\min_{\vartheta}\sum_{i=1}^{n}\Big\{\rho_\tau\Big(W_i-\vartheta^{\mathrm{T}}\widetilde{U}_i-\frac{\widehat{\xi}_i}{\sqrt{nh_1}}+\varepsilon_i\Big)-\rho_\tau\Big(W_i-\frac{\widehat{\xi}_i}{\sqrt{nh_i}}+\varepsilon_i\Big)\Big\}K_3\Big(\frac{U_i-u_0}{h_3}\Big),$$

where $\widehat{\xi}_i=\sqrt{nh_1}[\widehat{a}_0(X_i)-a_0(X_i)]$. Obviously,

$$\widehat{\vartheta}_1=\sqrt{nh_3}(\widehat{b}_0-a(u_0))=\sqrt{nh_3}(\widehat{a}^{\mathrm{e}}(u_0)-a(u_0)).\tag{4.8}$$

Thus, to prove Theorem 2.4, it suffices to show that $\widehat{\vartheta}_1$ is asymptotic normal. Denote by $\vartheta^*=\frac{1}{2}\sqrt{n}h_3^{\frac{5}{2}}(\mu_2^{(3)}a''(u_0)^{\mathrm{T}},0^{\mathrm{T}})^{\mathrm{T}}$ and

$$\widetilde{\vartheta}=\vartheta^*+\frac{\Sigma^*(u_0)^{-1}}{f_u(u_0)}\sum_{i=1}^{n}\psi_\tau(\varepsilon_i)\widetilde{U}_iK_3\Big(\frac{U_i-u_0}{h_3}\Big),\tag{4.9}$$

where $\Sigma^*(u_0) = \mathrm{diag}(\Sigma_0^*(u_0), \Sigma_0^*(u_0)\mu_2^{(3)})$. So to finish the proof of Theorem 2.4, it suffices to prove that for any $\delta > 0$,

$$P\{\|\widehat{\vartheta} - \widetilde{\vartheta}\| < \delta\} \to 1. \tag{4.10}$$

Let

$$G_n(\vartheta, \xi) = \sum_{i=1}^n \Big\{\rho_\tau\Big(W_i - \vartheta^{\mathrm{T}}\widetilde{U}_i - \frac{\xi}{\sqrt{nh_1}} + \varepsilon_i\Big) - \rho_\tau\Big(W_i - \frac{\xi}{\sqrt{nh_1}} + \varepsilon_i\Big)\Big\} K_3\Big(\frac{U_i - u_0}{h_3}\Big).$$

Write $\widehat{\xi} = \sqrt{nh_1}[\widehat{a}_0(x_0) - a_0(x_0)]$. Then, $\widehat{\xi} = O_p(1)$ by Theorem 2.2. Thus, by using the properties of $\rho_\tau(\,\cdot\,)$, it suffices to show that for any sufficient large $L > 0, L^* > 0$,

$$P\Big(\inf_{\substack{\|\vartheta - \widetilde{\vartheta}\| = \delta \\ \|\xi\| \le L}} [G_n(\vartheta, \xi) - G_n(\widetilde{\vartheta}, \xi)] > 0\} \cap \{\|\widetilde{\vartheta}\| \le L^*\}\Big) \to 1. \tag{4.11}$$

Let

$$\begin{aligned} &\Phi_n(\vartheta, \xi) = E[G_n(\vartheta, \xi)], \\ &\Psi_n(\vartheta, \xi) = G_n(\vartheta, \xi) - \Phi_n(\vartheta, \xi) + \sum_{i=1}^n \psi_\tau(\varepsilon_i)\vartheta^{\mathrm{T}}\widetilde{U}_i K_3\Big(\frac{U_i - u_0}{h_3}\Big). \end{aligned} \tag{4.12}$$

By assumption (C5), it can be obtained that

$$\begin{aligned} E[G_n(\vartheta, \xi)] &= \sum_{i=1}^n E\Big\{\int_{W_i - \frac{\xi}{\sqrt{nh_1}}}^{W_i - \vartheta^{\mathrm{T}}\widetilde{U}_i - \frac{\xi}{\sqrt{nh_1}}} [\tau - F_{\varepsilon\,|\,U_i, Z_i}(-s)]\mathrm{d}s K_3\Big(\frac{U_i - u_0}{h_3}\Big)\Big\} \\ &= \frac{1}{2}\sum_{i=1}^n E\Big\{f_{\varepsilon\,|\,U_i, Z_i}(0)\Big[(\vartheta^{\mathrm{T}}\widetilde{U}_i)^2 - \vartheta^{\mathrm{T}}\widetilde{U}_i W_i + \vartheta^{\mathrm{T}}\widetilde{U}_i\frac{\xi}{\sqrt{nh_1}}\Big]K_3\Big(\frac{U_i - u_0}{h_3}\Big)\Big\} + o(1) \\ &= \frac{1}{2}f_u(u_0)[\vartheta^{\mathrm{T}}\Sigma^*(u_0)\vartheta - \sqrt{n}h_3^{\frac{5}{2}}\mu_2^{(3)}\vartheta^{\mathrm{T}}(\Sigma_0^*(u_0)^{\mathrm{T}}, 0)^{\mathrm{T}}a''(u_0)] + o(1). \end{aligned} \tag{4.13}$$

From (4.9), we have

$$\sum_{i=1}^n \psi_\tau(\varepsilon_i)\vartheta^{\mathrm{T}}\widetilde{U}_i K_3\Big(\frac{U_i - u_0}{h_3}\Big) = f_u(u_0)\Big[\vartheta^{\mathrm{T}}\Sigma^*(u_0)\widetilde{\vartheta} - \frac{1}{2}\sqrt{n}h_3^{\frac{5}{2}}\mu_2^{(3)}\vartheta^{\mathrm{T}}(\Sigma_0^*(u_0)^{\mathrm{T}}, 0)^{\mathrm{T}}a''(u_0)\Big].$$

Since $2\vartheta^{\mathrm{T}}\Sigma^*(u_0)\widetilde{\vartheta} = \vartheta^{\mathrm{T}}\Sigma^*(u_0)\vartheta + \widetilde{\vartheta}^{\mathrm{T}}\Sigma^*(u_0)\widetilde{\vartheta} - (\vartheta - \widetilde{\vartheta})^{\mathrm{T}}\Sigma^*(u_0)(\vartheta - \widetilde{\vartheta})$, and by (4.12)–(4.13) and the above, we have

$$G_n(\vartheta, \xi) = \frac{1}{2}f_u(u_0)[(\vartheta - \widetilde{\vartheta})^{\mathrm{T}}\Sigma^*(u_0)(\vartheta - \widetilde{\vartheta}) - \widetilde{\vartheta}^{\mathrm{T}}\Sigma^*(u_0)\widetilde{\vartheta}] + \Psi_n(\vartheta, \xi) + o_p(1). \tag{4.14}$$

By using the above, for $\widetilde{\vartheta}$ satisfying that $\|\widetilde{\vartheta}\| \le L^*$, we have

$$G_n(\widetilde{\vartheta}, \xi) = -\frac{1}{2}f_u(u_0)\widetilde{\vartheta}^{\mathrm{T}}\Sigma^*(u_0)\widetilde{\vartheta} + \Psi_n(\widetilde{\vartheta}, \xi) + o_p(1). \tag{4.15}$$

Note that $\|\vartheta - \widetilde{\vartheta}\| = \delta$, and from (4.14) and (4.15), it follows that

$$G_n(\vartheta, \xi) - G_n(\widetilde{\vartheta}, \xi) \ge \frac{\delta^2}{2}f_u(u_0)\lambda_{\min}(\Sigma^*(u_0)) - 2\sup_{\substack{\|\xi\| \le L \\ \|\vartheta\| \le L^* + \delta}} \|\Psi_n(\vartheta, \xi)\| + o_p(1), \tag{4.16}$$

where $\lambda_{\min}(\Sigma^*(u_0))$ is the smallest eigenvalue of $\Sigma^*(u_0)$. Under the assumptions in Theorem 2.4, by using the argument similar to [2, Theorem 1], we can obtain

$$\sup_{\substack{\|\xi\|\le L \\ \|\vartheta\|\le L^*+\delta}} \|\Psi_n(\vartheta,\xi)\| = o_p(1).$$

This, together with (4.16) and assumption (C6), proves that (4.11) holds, and consequently (4.10) holds. The proof of Theorem 2.4 is completed.

Proof of Theorem 2.5 By using the similar argument as in Theorem 2.4, we can prove Theorem 2.5. Here, we omit the process of the proof.

Acknowledgements The authors sincerely thank Professor Lu and the referees for their valuable comments and helpful suggestions, which greatly improved the quality of the paper.

References

[1] Cai, Z., Fan, J. and Yao, Q., Functional-coefficient regression models for nonlinear time series, *J. Amer. Statist. Assoc.*, **95**, 2000, 941–956.

[2] Cai, Z. and Xu, X., Nonparametric quantile estimations for dynamic smooth coefficient models, *J. Amer. Statist. Assoc.*, **104**, 2009, 371–383.

[3] Engle, R., Granger, C., Rice, J., et al., Semiparametric estimates of the relation between weather and electricity sales, *J. Amer. Statist. Assoc.*, **81**, 1986, 310–320.

[4] Fan, J. and Gijbels, I., Local Polynomial Modelling and Its Applications, Chapment & Hall, London, 1996.

[5] Härdle, W., Liang, H. and Gao, J., Partially Linear Models, Springer-Verlag, Berlin, 2000.

[6] Hastie, T. and Tibshirani, R., Varying-coefficient models, *Roy. Statist. Soc.*, **55**, 1993, 757–796.

[7] Honda, T., Quantile regression in varying coefficient models, *J. Statist. Plann. Inference*, **121**, 2004, 113–125.

[8] Kim, M., Quantile Regression with varying coefficients, *Ann. Statist.*, **35**, 2007, 92–108.

[9] Koenker, R. and Zhao, Q., Conditional quantile estimation and inference for ARCH models, *Econometric Theory*, **12**, 1996, 793–813.

[10] Linton, O. B., Efficient estimation of additive nonparametric regression models, *Biometrika*, **82**, 1997, 93–100.

[11] Mu, Y. and Wei, Y., A dynamic quantile regression transformation model for longitudinal data, *Statist. Sin.*, **19**, 2009, 1137–1153.

[12] Wang, H., Zhu, Z. and Zhou, J., Quantile regression in partially linear varying coefficient models, *Ann. Statist.*, **37**, 2009, 3841-3866.

[13] Wong, H., Zhang, R., Ip, W. C., et al., Functional-coefficient partially regression model, *J. Multivariate Anal.*, **99**, 2008, 278–305.

[14] Zhang, W. and Lee, S., Local polynomial fitting in semivarying coefficient models, *J. Multivariate Anal.*, **82**, 2003, 166–188.