

Semi-functional Partial Linear Quantile Regression *

Hui Ding^{1,2}, Zhiping Lu¹, Jian Zhang³, Riquan Zhang^{1†}

1. School of Statistics, East China Normal University, Shanghai, 200241, P. R. China
2. School of Economics, Nanjing University of Finance and Economics, Nanjing, 210023, P. R. China
3. School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent CT2 7NF, UK

July 19, 2018

Abstract

Semi-functional partial linear model is a flexible model in which a scalar response is related to both functional covariate and scalar covariates. We propose a quantile estimation of this model as an alternative to the least square approach. We also extend the proposed method to kNN quantile method. Under some regular conditions, we establish the asymptotic normality of quantile estimators of regression coefficient. We also derive the rates of convergence of nonparametric function. Finite-sample performance of our estimation is compared with least square approach via a Monte Carlo simulation study. The simulation results indicate that our method is much more robust than the least square method. A real data example about spectrometric data is used to illustrate that our model and approach are promising.

Key words and phrases: *Functional data analysis; Partial linear; Quantile regression.*

AMS 2000 Subject Classifications: *Primary 62G05, Secondary 62G20.*

Short title: *Semi-functional Partial Linear Quantile Regression*

1 Introduction

Over the last two decades, technological progress in many subject areas have produced a large number of continuous data with curves or images as the units of observation. Functional data analysis (FDA) encompasses the statistical methodology for such data and has been prevailed. See Müller (2005); Cuevas (2014); Morris (2015); Wang et al. (2016); Goia and Vieu (2016) for systematic reviews on this subject. The recent monographs by Horváth and Kokoszka (2012); Hsing and Eubank (2015) offer some mathematical theories of functional data. As the important tool of FDA, functional regression aim to model the relationship between functional (scalar) response and functional (scalar) covariates. Researchers are increasingly interested in functional regression models. See Greven and Scheipl (2017) for a short survey on this field. It is noteworthy that semiparametric functional regression models offer a well-balanced mixture of parametric models and nonparametric models. Semiparametric functional regression models keep flexibility of parametric regression models and overcome sensitivity to dimensional effects of nonparametric approaches. See Goia and Vieu (2014) for a short survey. Semi-functional partial linear regression model is an important semiparametric

*The research was supported in part by National Natural Science Foundation of China (11571112,11701360), Program of Shanghai Subject Chief Scientist (14XD1401600), the 111 Project of China (B14019), Research Innovation Program for ECNU Graduates (ykc17083), the Natural Science Foundation of the Anhui Provincial Department of Education (KJ2017A424,KJ2018A0424) and the Anhui Province Natural Science Foundation(1508085QA14).

†Correspondence address: School of Statistics, East China Normal University, Shanghai, 200241, P. R. China. E-mail: zhangriquan@163.com.

functional regression model. It can be expressed as $Y = m(X) + \mathbf{Z}^\top \boldsymbol{\beta} + \varepsilon$, where X is functional covariate that taking values in semi-metric space \mathcal{F} , $m : \mathcal{F} \rightarrow \mathcal{R}$ is a unknown smooth function, \mathbf{Z} is p -dimensional random vector of scalar covariate, $\boldsymbol{\beta}$ is unknown coefficient of scalar covariate, ε is a random error. The model has been widely used in many fields. Aneiros-Pérez and Vieu (2006) proposed profile least square estimation method and derived the asymptotic performances of proposed estimators. Aneiros-Pérez and Vieu (2008) extended the model to time series area. Ling et al. (2017) proposed a k-nearest-neighbours (kNN) procedure and derived the asymptotic performances of kNN estimators. Aneiros et al. (2015) extend the model to high-dimensional framework. They proposed penalized least-squares method to study the problem of variable selection and derived an oracle property.

The above-mentioned references are all focused on mean regression. It is known that mean regression is sensitive to outliers. Quantile regression is usually recognized as an alternative to mean regression. Quantile regression is more robust than mean regression. There is few literatures on quantile-regression-based estimation procedures in the functional regression model. Cardot et al. (2005) proposed a spline estimator for functional linear quantile regression model. Chen and Müller (2012) studied a estimation method for conditional quantile analysis in the generalized functional regression framework. Kato et al. (2012) studied quantile estimation in functional linear quantile regression model. In this paper, we study quantile regression of semi-functional partial linear model. To the best of our knowledge, this method has not been researched in the scientific literature. Since the model is flexible in practice, quantile regression method is urgently needed, which motivates us to investigate quantile regression of estimation of the model.

In this paper, we use quantile regression method to estimate the nonparametric function and regression coefficient of the model. We also extend the proposed method to kNN quantile method. Under some regular conditions, we establish the asymptotic normality of estimators of regression coefficient and derive the rates of convergence of nonparametric function. A Monte Carlo simulation and an application to spectrometric data show the advantages of our proposed method.

The article is organized as follows. Section 2 describes our model and our estimation method. Section 3 and 4 present asymptotic properties and finite sample performance of the proposed estimators respectively. Section 5 provides an application to spectrometric data. Concluding remarks are provided in Section 6. Technical proofs are given in an Appendix.

2 Model and estimation

2.1 Model

Given quantile level $\tau \in (0, 1)$, we consider the following semi-functional partial linear quantile regression model

$$Y = m_\tau(X) + \mathbf{Z}^\top \boldsymbol{\beta}_\tau + \varepsilon_\tau, \quad (1)$$

where X is functional covariate that taking values in semi-metric space \mathcal{F} , and we denote the associated semi-metric by $d(\cdot, \cdot)$, $m_\tau : \mathcal{F} \rightarrow \mathcal{R}$ is a unknown smooth function, $\mathbf{Z} = (Z_1, \dots, Z_p)^\top$ are p -dimensional random vector of scalar covariates, $\boldsymbol{\beta}_\tau = (\beta_{1\tau}, \dots, \beta_{p\tau})^\top$ are unknown coefficients of scalar covariates, ε_τ is a random error whose τ th quantile conditional on (\mathbf{Z}, X) being zero.

2.2 Estimation

Suppose that $\{(Y_i, X_i, \mathbf{Z}_i), i = 1, \dots, n\}$ is a random sample generated from model (1). We estimate coefficients $\boldsymbol{\beta}_\tau$ and function $m_\tau(\cdot)$ in model (1), by minimizing the following quantile loss function

$$\sum_{i=1}^n \rho_\tau \left(Y_i - m_\tau(X_i) - \mathbf{Z}_i^\top \boldsymbol{\beta}_\tau \right), \quad (2)$$

where $\rho_\tau(s) = s\{\tau - I(s < 0)\}$.

Obviously, (2) contains both nonparametric and parametric component. And they can be estimated by different rates of convergence, so we use three-stage procedure. In the first stage, we apply the local constant weighted quantile smoothing technique to get an initial estimators. That is, we obtain an initial estimators of $m_\tau(X_j)$ and β_τ by minimizing the following local weighted quantile loss function

$$\sum_{i:i \neq j} \rho_\tau \left(Y_i - a_{\tau j} - \mathbf{Z}_i^\top \beta_\tau \right) K_{h_0}(d(X_i, X_j)), \quad (3)$$

where $K_{h_0}(\cdot) = K(\cdot/h_0)/h_0$ and $K(\cdot)$ is a kernel function and h_0 is a bandwidth. For convenience, we denote the initial estimators of $a_{\tau j}, \beta_\tau$ by $\tilde{a}_{\tau j}, \tilde{\beta}_\tau$. In the second stage, we further improve the efficiency of $\tilde{\beta}_\tau$. Specifically, we derive the final estimator of β_τ by minimizing the following quantile loss function

$$\sum_{j=1}^n \rho_\tau \left(Y_j - \tilde{a}_{\tau j} - \mathbf{Z}_j^\top \beta_\tau \right). \quad (4)$$

Denote the final estimator of β_τ by $\hat{\beta}_\tau$. In the third stage, we obtain the final estimator of $m_\tau(x)$. More concretely, we obtain the final estimator of $m_\tau(x)$ by minimizing the the following local weighted quantile loss function

$$\sum_{i=1}^n \rho_\tau \left(Y_i - a_\tau - \mathbf{Z}_i^\top \hat{\beta}_\tau \right) K_h(d(X_i, x)). \quad (5)$$

Denote the final estimators of a_τ by \hat{a}_τ . Evidently, \hat{a}_τ are the final estimator of $m_\tau(x)$.

Remark 1 The estimation procedure proposed above get the estimators of the $m_\tau(x)$ and β_τ by minimizing quantile loss function. In next section, we show that minimizing (3) can help us to obtain initial estimators $\tilde{\beta}_\tau$ for β_τ , and get a improved root-n consistent estimator $\hat{\beta}_\tau$ eventually if we continue to minimize (4). We also show that \hat{a}_τ is the strong consistent estimator of $m_\tau(x)$ and derive the rate of convergence of estimator.

2.3 Tuning parameters selection

To implement our estimation method, we need to know how to choose semi-metric $d(\cdot, \cdot)$ and bandwidths h_0 and h . We considered (as recommended in Ferraty and Vieu (2006)) a class of semi-metrics based on derivatives $\{d_q^{[a,b]}(\cdot, \cdot), q = 0, 1\}$, where

$$d_q^{[a,b]}(x_i, x_j) = \left(\int_a^b \left(x_i^{(q)}(t) - x_j^{(q)}(t) \right)^2 dt \right)^{1/2}.$$

We will use leave-one-out cross-validation to select (q, h_0) and (q, h) . The observation j is removed to compute $\tilde{\beta}_{\tau, q, h_0}^{(-j)}$ and $\tilde{a}_{\tau, j, q, h_0}^{(-j)}$, where $\tilde{\beta}_{\tau, q, h_0}^{(-j)}$ and $\tilde{a}_{\tau, j, q, h_0}^{(-j)}$ denoting the initial leave-one-out estimator of β_τ and $m_\tau(X_j)$. Thus, for given (q, h_0) , we compute

$$CV_1(q, h_0) = \frac{1}{n} \sum_{j=1}^n \rho_\tau \left(Y_j - \tilde{a}_{\tau, j, q, h_0}^{(-j)} - \mathbf{Z}_j^\top \tilde{\beta}_{\tau, q, h_0}^{(-j)} \right).$$

Then, (q, h_0) is obtained by $(q, h_0) = \arg \min_{q, h_0} CV_1(q, h_0)$. Once (q, h_0) is obtained, $\tilde{a}_{\tau j}$ and $\tilde{\beta}_\tau$ can be computed using this tuning parameters. Then, $\hat{\beta}_\tau$ can be obtained. To provide a selector of (q, h) , a second leave-one-out cross-validation can be used. More precisely, (q, h) is defined as $(q, h) = \arg \min_{q, h} CV_2(q, h)$. where

$$CV_2(q, h) = \frac{1}{n} \sum_{j=1}^n \rho_\tau \left(Y_j - \hat{a}_{\tau, j, q, h}^{(-j)} - \mathbf{Z}_j^\top \hat{\beta}_\tau \right),$$

$\hat{a}_{\tau, j, q, h}^{(-j)}$ denote that the final estimator of $m_\tau(X_j)$ computed without observation j and using tuning parameters (q, h) .

2.4 Extension to kNN quantile method

The methodology and tuning parameters selection of the previous sections are based on the fixed bandwidth parameter, which ignore the local denseness/sparseness in practice. In this section, we propose kNN quantile method to address this problem. It is noteworthy that our proposed estimation procedure in Section 2.2 can be easily extended to kNN quantile method. A key step is to replace fixed bandwidth h_0 and h with local adaptive neighbourhood size, based on the distance of a point from its neighbours. Let $\mathcal{B}(x, h)$ be the open ball centered at $x \in \mathcal{F}$ and of radius $h > 0$. We also use three-stage procedure similar to that used in Section 2.2. We only need to replace (3) and (5) with

$$\sum_{i:i \neq j} \rho_\tau \left(Y_i - a_{\tau j} - \mathbf{Z}_i^\top \boldsymbol{\beta}_\tau \right) K \left(\frac{d(X_i, X_j)}{H_{k_0}(X_j)} \right) \quad \text{and} \quad \sum_{i=1}^n \rho_\tau \left(Y_i - a_\tau - \mathbf{Z}_i^\top \hat{\boldsymbol{\beta}}_\tau \right) K \left(\frac{d(X_i, x)}{H_k(x)} \right)$$

respectively, where $H_k(x) = \min\{h > 0 : \sum_{i=1}^n I_{\mathcal{B}(x, h)}(X_i) = k\}$. By using leave-one-out cross-validation method similar to that used in Section 2.3, we can choose optimal tuning parameters (q, k_0) and (q, k) .

Remark 1 kNN quantile procedure has two major advantages over quantile procedure in Section 2.2. First, kNN quantile procedure includes local adaptive smoothing technology, which is a key point in infinite dimensional analysis for taking into account local denseness/sparseness structures of the data. Second, kNN quantile procedure uses a discrete smoothing parameter k which is much more easier to choose in practice than the continuous bandwidth h .

3 Asymptotic properties

In this section we only establish asymptotic properties of the quantile estimators proposed in Section 2.2. Let $f_\tau(\cdot | \mathbf{z}, x)$ and $F_\tau(\cdot | \mathbf{z}, x)$ denote the density function and cumulative distribution function of the error ε_τ condition on $(\mathbf{Z}, X) = (\mathbf{z}, x)$ respectively. Let $\mathbf{G}(x) = \mathbb{E}\{f_\tau(0 | \mathbf{Z}, X)(1, \mathbf{Z}^\top)^\top(1, \mathbf{Z}^\top) | X = x\}$, $\mathbf{A}_\tau(x, \mathbf{z}) = \mathbb{E}\{f_\tau(0 | \mathbf{Z}, X) \mathbf{Z} (1, \mathbf{0}^\top) | X = x\} \{\mathbf{G}(x)\}^{-1} \mathbf{z}$, $\mathbf{H}(x) = \mathbb{E}\{f_\tau(0 | \mathbf{Z}, X) \mathbf{Z} | X = x\}$, $\mathbf{M}_\tau = \mathbb{E}\{f_\tau(0 | \mathbf{Z}, X) \mathbf{Z} \mathbf{Z}^\top\}$ and $\mathbf{N}_\tau = \mathbb{E}\{[\mathbf{Z} - \mathbf{A}_\tau(X, \mathbf{Z})][\mathbf{Z} - \mathbf{A}_\tau(X, \mathbf{Z})]^\top\}$. Let $S_{\mathcal{F}}$ be a given compact subset of \mathcal{F} . We assume that X is valued in $S_{\mathcal{F}}$. For any $\epsilon > 0$, $\psi_{S_{\mathcal{F}}} = \log(N_\epsilon(S_{\mathcal{F}}))$ denotes the Kolmogorov's ϵ -entropy of $S_{\mathcal{F}}$. ($N_\epsilon(S_{\mathcal{F}})$ is the minimal number of open balls in $S_{\mathcal{F}}$ of radius ϵ which is necessary to cover $S_{\mathcal{F}}$). The following conditions are needed:

- (C1) For any $x \in S_{\mathcal{F}}$, assume that there exist positive constants c_1 and c_2 and a function $\phi(h)$ on $(0, \infty)$ such that $0 < c_1 \phi(h) \leq P(X \in \mathcal{B}(x, h)) \leq c_2 \phi(h) < \infty$.
- (C2) We assume that m_τ is smooth, in the sense that there exist constants $c_3 > 0$ and $\alpha > 0$ such that for any $u, v \in S_{\mathcal{F}}$, $|m_\tau(u) - m_\tau(v)| \leq c_3(d(u, v))^\alpha$.
- (C3) The kernel function $K(\cdot)$ satisfies:
 - (C3.i) $K(\cdot)$ is a bounded nonnegative function with support $[0, 1]$ and satisfies a Lipschitz condition on $[0, 1)$.
 - (C3.ii) If $K(1) = 0$, the kernel function $K(\cdot)$ has to satisfy an addition condition: its derivative $K'(\cdot)$ exists and there exists two constants c_3 and c_4 such that $-\infty < c_3 \leq K'(u) \leq c_4 < 0$.
- (C4) The function ϕ satisfies:
 - (C4.i) There exists two positive constants c_5 and η_0 such that $\phi'(\eta) < c_5$ for any $\eta < \eta_0$.
 - (C4.ii) If $K(1) = 0$, the function $\phi(\cdot)$ has to satisfy an addition condition: there exists two positive constants c_5 and η_0 such that $\int_0^\eta \psi(u) du > c_5 \eta \psi(\eta)$ for any $0 < \eta < \eta_0$.
- (C5) Kolmogorov's ϵ -entropy of $S_{\mathcal{F}}$ satisfies: for n large enough, $(\log n)^2/n\phi(h) < \psi_{S_{\mathcal{F}}}(\frac{\log n}{n}) < n\phi(h)/\log n$.

(C6) There exist constant M such that $\forall m \geq 2$, $E \left\{ |f_\tau(0|\mathbf{Z}, X) \mathbf{Z} \mathbf{Z}^\top|^m | X = x \right\} < \sigma_m(x) \leq M < \infty$ with $\sigma_m(\cdot)$ continuous on $S_{\mathcal{F}}$.

(C7) $F_\tau(0|\mathbf{z}, x) = \tau$ for all (\mathbf{z}, x) , $f_\tau(\cdot|\mathbf{z}, x)$ is bounded away from zero and has a continuous and uniformly bounded derivative.

(C8) \mathbf{Z} has bounded support.

(C9) $\mathbf{G}(x)$ and $\mathbf{H}(x)$ are continuous on $S_{\mathcal{F}}$, $\mathbf{G}(x)$ is nonsingular on $S_{\mathcal{F}}$.

Remark 1 Condition C1-C9 are not the weakest possible conditions, but they are imposed to facilitate the proof of Theorem. Conditions C1-C6 are required in the context of semi-functional partial linear model (see Ling et al., 2017). They are also a direct extension of Aneiros-Pérez and Vieu (2006). Specifically, it increases additional topological conditions to obtain uniform convergence results (see Ferraty et al., 2010; Kudraszow and Vieu, 2013). Conditions C7-C9 are needed for quantile regression. They are quite usual in nonfunctional quantile regression (see Kai et al., 2011). Condition C1-C6 and C9 are standard in the nonparametric functional regression setting and extend what is usually assumed in the classical p -dimensional nonparametric literature (see Ferraty and Vieu, 2006).

Theorem 1. *Under the regularity conditions C1-C9, if in addition h_0 satisfies C1 and C5, $h_0 \rightarrow 0$, $\frac{nh_0^{4\alpha}}{\psi_{S_{\mathcal{F}}}(\frac{\log n}{n})} \rightarrow 0$ and $\frac{\psi_{S_{\mathcal{F}}}(\frac{\log n}{n})}{n\phi^2(h_0)} \rightarrow 0$ as $n \rightarrow \infty$, then*

$$\sqrt{n}(\hat{\beta}_\tau - \beta_\tau) \xrightarrow{\mathcal{L}} N(\mathbf{0}, \tau(1-\tau)\mathbf{M}_\tau^{-1}\mathbf{N}_\tau\mathbf{M}_\tau^{-1}). \quad (6)$$

Theorem 2. *Suppose that the regularity conditions C1-C9 hold. If $h \rightarrow 0$ as $n \rightarrow \infty$, then*

$$\sup_{x \in S_{\mathcal{F}}} |\hat{m}_\tau(x) - m_\tau(x)| = O_p \left(h^\alpha + \sqrt{\frac{\psi_{S_{\mathcal{F}}}(\frac{\log n}{n})}{n\phi(h)}} \right). \quad (7)$$

Remark 2 The condition about kernel bandwidth h_0 in Theorem 1 ensures that we can derive the root- n consistency and asymptotic normality for estimator $\hat{\beta}_\tau$. Under the condition about kernel bandwidth h in Theorem 2, we can get the rate of uniformly convergence of nonparametric function.

4 Simulation studies

In this section, we implement a simulation study to investigate the performance of the proposed estimation methods. The data sets are generated from the following model:

$$Y_i = Z_{1i}\beta_1 + Z_{2i}\beta_2 + m(X_i) + 0.3\varepsilon_i, \quad i = 1, 2, \dots, n.$$

For the functional nonparametric component, we take the same form as Model M1 of Aneiros-Pérez and Vieu (2011), that is, The functional data $X_i(t) = a_i(t - 0.5)^2 + b_i(t \in [0, 1])$ with a_i and b_i are independent and identically distributed uniform random variables $U(0, 1)$ and $U(-0.5, 0.5)$, respectively. $g(X_i) = \text{sign}(X_i'(1) - X_i'(0)) \sqrt{3 \int_0^1 (X_i'(t))^2 dt}$ and $m(X_i) = \exp(-8g(X_i)) - \exp(-12g(X_i))$, where $\text{sign}(a)$ denotes the sign of a . For the linear component, we let Z_{1i} and Z_{2i} are independent and identically distributed normal random variables $N(0, 1)$. The coefficients of covariates are $\beta_1 = -1$ and $\beta_2 = 3$. Furthermore, X_i , Z_{1i} and Z_{2i} are independent.

In our simulation, we consider two cases for error terms ε : $N(0, \sigma^2)$ and standard Cauchy, where $\sigma = \max_X(m(X)) - \min_X(m(X))$. We also consider two choices for the sample sizes $n = 100$ and 300 . Each

$X_i(t)$ is observed at 100 equally space points on $[0, 1]$. We use the cross-validation procedure as described in section 2.3 to select optimal semi-metric and bandwidths. We also use kNN quantile method. We apply the cross-validation procedure as described in section 2.4 to select optimal semi-metric and neighbourhood size. The Epanechnikov kernel is used in both the quantile regression procedure and kNN quantile regression procedure. In order to evaluate the performance of estimators of different method, we compare the profile least squares (PLS) method (see Aneiros-Pérez and Vieu, 2006), our quantile regression (QR) method and kNN quantile regression (kNNQR) method. We focus on $\tau = 0.25, 0.5$ and 0.75 in quantile regression. All simulations are replicated for 1000 times.

Performance of the estimators of coefficient functions β_1 and β_2 are assessed using the bias and standard deviation of estimators. While performance of estimator of functional nonparametric function $m(x)$ is assessed using the square root of average squared errors (RASE) defined as

$$\text{RASE}\{\hat{m}(x)\} = \left\{ \frac{1}{100} \sum_{i=1}^{100} [\hat{m}(x_i) - m(x_i)]^2 \right\}^{\frac{1}{2}},$$

where $x_i, i = 1, 2, \dots, 100$ are new random samples of size 100 which generated from $X_i(t)$.

Table 1-2 list RASEs of $\hat{m}(x)$ and biases and standard deviations of $\hat{\beta}_1$ and $\hat{\beta}_2$ under different error terms. There is a general tendency for RASEs of $\hat{m}(x)$ and biases and standard deviations of $\hat{\beta}_1$ and $\hat{\beta}_2$ to decrease as sample sizes increases. From Table 1, we can see that PLS estimators, QR estimators and kNNQR estimators have small RASEs and biases under normal error terms. QR estimators and kNNQR estimators are slightly worse than PLS estimators as expected. When the error term follows heavy-tailed standard Cauchy distribution, Table 2 illustrate that PLS estimators have very large RASEs and biases and standard deviations while QR estimators and kNNQR estimators have reasonably small RASEs and biases and standard deviations. Note that this conclusion agrees with the known fact that PLS fails when the error variance is infinite. Overall, Our proposed estimation methods shows better performance even with infinite variance errors. Focusing on the QR and kNNQR estimation of the nonparametric component $m(\cdot)$, we can see that kNNQR estimators are better than QR estimators. This is because kNNQR method is similar to QR method with a local adaptive variable bandwidth.

Table 1: RASEs and biases with standard deviations(in parentheses) with normal distribution error $N(0, \sigma^2)$

n	Method	$\hat{m}(x)$	$\hat{\beta}_1$	$\hat{\beta}_2$
100	PLS	0.0017(0.0003)	0.0029(0.0062)	-0.0023(0.0050)
	QR(0.25)	0.0045(0.0008)	0.0041(0.0068)	-0.0018(0.0067)
	QR(0.50)	0.0022(0.0004)	0.0028(0.0083)	-0.0026(0.0087)
	QR(0.75)	0.0038(0.0008)	0.0039(0.0104)	-0.0043(0.0082)
	kNNQR(0.25)	0.0040(0.0013)	0.0034(0.0069)	-0.0018(0.0050)
	kNNQR(0.50)	0.0016(0.0004)	0.0036(0.0102)	-0.0026(0.0071)
	kNNQR(0.75)	0.0034(0.0010)	0.0029(0.0091)	-0.0068(0.0081)
300	PLS	0.0012(0.0001)	-0.0001(0.0036)	-0.0003(0.0028)
	QR(0.25)	0.0041(0.0004)	0.0004(0.0047)	0.0001(0.0048)
	QR(0.50)	0.0018(0.0002)	-0.0003(0.0056)	0.0003(0.0045)
	QR(0.75)	0.0032(0.0005)	0.0001(0.0045)	-0.0018(0.0041)
	kNNQR(0.25)	0.0038(0.0011)	0.0008(0.0049)	0.0006(0.0044)
	kNNQR(0.50)	0.0015(0.0002)	0.0001(0.0057)	-0.0005(0.0048)
	kNNQR(0.75)	0.0030(0.0010)	0.0010(0.0065)	-0.0018(0.0040)

Table 2: RASEs and biases with standard deviations(in parentheses) with standard Cauchy distribution error

n	Method	$\hat{m}(x)$	$\hat{\beta}_1$	$\hat{\beta}_2$
100	PLS	32335.9(686275.9)	3.1875(80.3151)	-0.9969(59.7885)
	QR(0.25)	0.1003(0.0366)	-0.0018(0.0823)	0.0017(0.0680)
	QR(0.50)	0.0075(0.0071)	-0.0031(0.0543)	0.0085(0.0455)
	QR(0.75)	0.1168(0.0529)	0.0191(0.0698)	0.0108(0.0999)
	kNNQR(0.25)	0.0439(0.0432)	-0.0067(0.0769)	0.0017(0.0720)
	kNNQR(0.50)	0.0042(0.0029)	0.0044(0.0450)	0.0077(0.0545)
	kNNQR(0.75)	0.0589(0.0496)	0.0306(0.0697)	-0.0102(0.0839)
300	PLS	20711.1(459763.3)	1.8931(58.3539)	0.2585(19.2496)
	QR(0.25)	0.1037(0.0357)	-0.0032(0.0479)	0.0010(0.0482)
	QR(0.50)	0.0050(0.0027)	0.0021(0.0270)	0.0129(0.0265)
	QR(0.75)	0.1002(0.0297)	0.0185(0.0544)	0.0080(0.0425)
	kNNQR(0.25)	0.0500(0.0305)	0.0024(0.0622)	-0.0092(0.0585)
	kNNQR(0.50)	0.0034(0.0012)	0.0028(0.0244)	0.0068(0.0193)
	kNNQR(0.75)	0.0682(0.0416)	0.0068(0.0567)	0.0094(0.0360)

5 A real example

In this section, we apply our proposed quantile estimation to spectrometric data which are available from <http://lib.stat.cmu.edu/datasets/teacator>. These data are obtained for 215 pieces of pure meat. Each data sample contains fat, protein, water contents of finely chopped meat and spectrometric curve. The three contents measured in percent, are determined by analytic chemistry. Spectrometric curve consist of 100 wavelengths absorbance spectrum records. Aneiros-Pérez and Vieu (2006) consider the following semi-functional partial linear model

$$Y = Z_1\beta_1 + Z_2\beta_2 + m(X) + \varepsilon.$$

where Y is the percentage of fat content, X is the spectrometric curve, Z_1 and Z_2 are the corresponding percentages of water content and protein content, respectively. Our aim is not to achieve a full case study but to provide a richer characterization of the data, allowing us to consider the impact of a covariate on the entire distribution of Y , not merely its conditional mean. Firstly, profile least squares method (see Aneiros-Pérez and Vieu, 2006) is used in the above model to analyse the normality of the residuals. The norm quantile-quantile of the residuals is shown in Figure 1(a), from which we can see apparently that the residuals cannot follow normal distribution. We also make a Shapiro-Wilk hypothesis test to judge the normality of the residuals. By Shapiro-Wilk test, we find that the p value is less than 3.63×10^{-19} . This reminds us further that the error cannot be normal, and the mean regression is unsuitable here.

Our quantile regression method and kNN quantile regression method when $\tau = 0.1, 0.2, \dots, 0.9$ are used here to analyse the impact of water content and protein content on the entire distribution of fat content. Both quantile regression estimators and kNN quantile regression estimators have similar performance. To save space, we only show the result of quantile regression estimators proposed in Section 2.2. Specifically, the kernel used in the real analysis is $K(u) = 0.75(1 - u^2)I_{[0,1]}(u)$. We also use the cross-validation procedure as described in section 2.3 to select the semi-metric and bandwidths. The estimators of coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ are shown in Table 3. We also construct pointwise median, 2.5% and 97.5% quantiles of coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ from $\tau = 0.1$ to 0.9 and show it in Figure 1(b) and (c). It is apparent that the impact of water content on fat content is negative across all quantiles, and that it decreases as the quantile increases. Table 3 and Figure 1(b) and (c) also indicate that the impact of protein content on fat content is negative across all quantiles, but it increases as the quantile increases. As it is known to all, the more fat within the finely chopped meat, the richer the taste, the higher the grade. In other words, the water content matters more for low-quality meat in terms of fat content, whereas the protein content matters more for high-quality meat in terms of fat content.

Table 3: Means and standard errors (in parentheses) of coefficients for different quantiles

τ	$\hat{\beta}_1$	$\hat{\beta}_2$
0.1	-0.8066(0.0125)	-2.1419(0.0436)
0.2	-0.7863(0.0093)	-1.9752(0.0333)
0.3	-0.9269(0.0051)	-1.3573(0.0184)
0.4	-0.9872(0.0038)	-1.1021(0.0137)
0.5	-1.0075(0.0030)	-1.0108(0.0106)
0.6	-1.0231(0.0021)	-0.9479(0.0076)
0.7	-1.0144(0.0022)	-0.9659(0.0079)
0.8	-1.0263(0.0022)	-0.9222(0.0074)
0.9	-1.0610(0.0023)	-0.7946(0.0075)

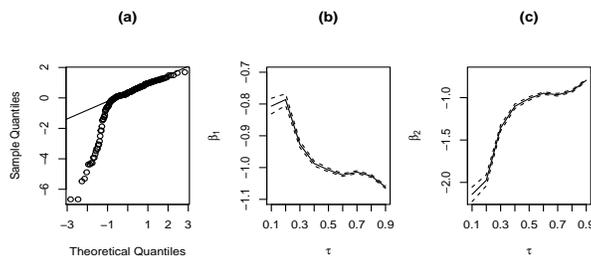


Figure 1: (a) QQ plot of the residual for profile least squares estimation method. (b) Pointwise median of $\hat{\beta}_1$ is shown as solid line from $\tau = 0.1$ to 0.9 . (c) Pointwise median of $\hat{\beta}_2$ is shown as solid line from $\tau = 0.1$ to 0.9 . Pointwise 2.5% and 97.5% quantiles are given as dotted lines in (b) and (c).

6 Concluding remarks

We have proposed a quantile estimation of semi-functional partial linear regression model. The estimators are obtained based on three-stage procedure. We also have extended the proposed method to kNN quantile method. We have established an asymptotic theory for the our proposed estimation. Finite-sample performances of our quantile estimation and kNN quantile estimation are compared with least square approach via a Monte Carlo simulation study. The simulation result indicate that our quantile method and kNN quantile method are much more robust than the least square method. A real data example about spectrometric data has been used to illustrate advantage of our quantile estimation method.

There are two interesting possible extensions of semi-functional partial linear quantile regression. First, we only propose the kNN quantile regression method. The kNN quantile regression method is based on a random neighbourhood. Hence, the theory of kNN quantile estimation is much more complicated and warrants further studies. Second, we note that our study focuses on finite-dimensional scalar covariates. However, in many fields of applications the number of scalar covariates is high. Thus, estimation and variable selection for extending the semi-functional partial linear quantile regression model with multiple scalar covariates is a challenge for future work.

7 Appendix

Proof. Let $r_i = I(\varepsilon_{i\tau} \leq 0) - \tau$ and $r_{ij} = I(\varepsilon_{i\tau} \leq \zeta_{ij}) - \tau$, where $\zeta_{ij} = m_\tau(X_j) - m_\tau(X_i)$. Denote $\tilde{\theta}_j = \sqrt{n\phi(h_0)}(\tilde{a}_{\tau j} - m(X_j), (\tilde{\beta}_\tau - \beta_\tau)^\top)^\top$, $\mathbf{Z}_i^* = \{1, \mathbf{Z}_i^\top\}^\top$, $\mathbf{Z}^* = \{1, \mathbf{Z}^\top\}^\top$ and $\eta_{ij} = \mathbf{Z}_i^{*\top} \tilde{\theta}_j / \sqrt{n\phi(h_0)}$.

For convenience, we use the symbol $\mathbf{A}_n = O_p(a_n)$ (or $o_p(a_n)$) to denote that the every element of matrix \mathbf{A}_n is $O_p(a_n)$ (or $o_p(a_n)$). For two sequences of positive numbers a_n and b_n , $a_n \lesssim b_n$ means a_n/b_n is uniformly bounded. $a_n \asymp b_n$ means $a_n \lesssim b_n$ and $b_n \lesssim a_n$.

Lemma A.1. *Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be independent and identically distributed random vectors, where X_i is functional random variable, Y_i is scalar random variable. We assume that X_i and $K(\cdot)$ satisfy conditions C1 and C3-C5. Assume that there exist constant M such that $\mathbb{E}(|Y^m||X = x) < \sigma_m(x) \leq M < \infty$ for $m \geq 2$, where $\sigma_m(\cdot)$ continuous on $S_{\mathcal{F}}$. Then*

$$\sup_{x \in S_{\mathcal{F}}} \left| \frac{1}{n\phi(h)} \sum_{i=1}^n [K(h^{-1}d(X_i, x))Y_i - \mathbb{E}K(h^{-1}d(X_i, x))Y_i] \right| = O_p \left(\sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right).$$

Proof. Let $K_i = K(h^{-1}d(X_i, x))$. By condition C1 and C3, if $K(1) > 0$, there exist positive constants c'_1 and c'_2 such that $c'_1\phi(h) < \mathbb{E}(K_i) < c'_2\phi(h)$ for any $x \in S_{\mathcal{F}}$. If $K(1) = 0$, by Lemma 4.4 of Ferraty and Vieu (2006), we can get the same results. Let $x_1, \dots, x_{N_\epsilon(S_{\mathcal{F}})}$ be an ϵ -net for $S_{\mathcal{F}}$ ($S_{\mathcal{F}} \subset \cup_{k=1}^{N_\epsilon(S_{\mathcal{F}})} \mathcal{B}(x_k, \epsilon)$) and $x_1, \dots, x_{N_\epsilon(S_{\mathcal{F}})} \in S_{\mathcal{F}}$. Set $k(x) = \arg \min_{k \in \{1, \dots, N_\epsilon(S_{\mathcal{F}})\}} d(x_k, x)$. Let $g(x) = \sum_{i=1}^n (K_i Y_i) / n\mathbb{E}(K_1)$, $G_1 = \sup_{x \in S_{\mathcal{F}}} |g(x) - g(x_{k(x)})|$, $G_2 = \sup_{x \in S_{\mathcal{F}}} |g(x_{k(x)}) - \mathbb{E}g(x_{k(x)})|$ and $G_3 = \sup_{x \in S_{\mathcal{F}}} |\mathbb{E}g(x_{k(x)}) - \mathbb{E}g(x)|$. By using analogous argument as for (13),(14) and (15) of Ferraty et al. (2010), we deduce that there exist constants $\omega > 1$, η and c' such that

$$\sup_n \mathbb{P} \left(\sup_{x \in S_{\mathcal{F}}} |g(x_{k(x)}) - \mathbb{E}g(x_{k(x)})| > \eta \sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right) \leq c' N_\epsilon(S_{\mathcal{F}})^{1-\omega} \rightarrow 0$$

and

$$G_1 = O_p \left(\sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right), G_3 = O_p \left(\sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right).$$

Thus, we have

$$\sup_{x \in S_{\mathcal{F}}} \left| \frac{1}{n\mathbb{E}(K_1)} \sum_{i=1}^n [K(h^{-1}d(X_i, x))Y_i - \mathbb{E}K(h^{-1}d(X_i, x))Y_i] \right| = O_p \left(\sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right).$$

Hence,

$$\begin{aligned} & \sup_{x \in S_{\mathcal{F}}} \left| \frac{1}{n\phi(h)} \sum_{i=1}^n [K(h^{-1}d(X_i, x))Y_i - \mathbb{E}K(h^{-1}d(X_i, x))Y_i] \right| \\ & \leq \sup_{x \in S_{\mathcal{F}}} \left| \frac{1}{n\mathbb{E}(K_1)} \sum_{i=1}^n [K(h^{-1}d(X_i, x))Y_i - \mathbb{E}K(h^{-1}d(X_i, x))Y_i] \right| \sup_{x \in S_{\mathcal{F}}} \left| \frac{\mathbb{E}(K_1)}{\phi(h)} \right| = O_p \left(\sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right). \end{aligned}$$

Lemma A.2. *Suppose that the regularity conditions C1-C9 hold. If $h_0 \rightarrow 0$ as $n \rightarrow \infty$, then*

$$\tilde{\theta}_j = - \left\{ \mathbb{E} \left\{ K(h_0^{-1}d(X, X_j)) \mathbb{E} \left[f_\tau(0|Z, X) \mathbf{Z}^* \mathbf{Z}^{*\top} | X \right] \right\} \right\}^{-1} \sqrt{\frac{\phi(h_0)}{n}} \sum_{i=1}^n \mathbf{Z}_i^* r_{ij} K(h_0^{-1}d(X_i, X_j)) + o_p(1).$$

Proof. Fix X_j and let $K_{ij} = K(h_0^{-1}d(X_i, X_j))$. Seen from (1), we deduce that

$$Y_i - \mathbf{Z}_i^\top \tilde{\boldsymbol{\beta}}_\tau - \tilde{a}_{\tau j} = \varepsilon_{i\tau} + \mathbf{Z}_i^\top (\boldsymbol{\beta}_\tau - \tilde{\boldsymbol{\beta}}_\tau) + m_\tau(X_i) - \tilde{a}_{\tau j} = \varepsilon_{i\tau} - \eta_{ij} - \zeta_{ij}.$$

Then, $\tilde{\boldsymbol{\theta}}_j$ is also the minimizer of $\ell_n^*(\boldsymbol{\theta}_j) = \sum_{i:i \neq j} \{\rho_\tau(\varepsilon_{i\tau} - \eta_{ij} - \zeta_{ij}) - \rho_\tau(\varepsilon_{i\tau} - \zeta_{ij})\} K_{h_0}(d(X_i, X_j))$. By the identity of Knight et al. (1998), we can get

$$\ell_n^*(\boldsymbol{\theta}_j) = \sum_{i:i \neq j} \left\{ \eta_{ij} r_{ij} + \int_0^{\eta_{ij}} \{I(\varepsilon_{i\tau} \leq \zeta_{ij} + t) - I(\varepsilon_{i\tau} \leq \zeta_{ij})\} dt \right\} K_{ij} = \mathbf{S}_j^{*\top} \boldsymbol{\theta}_j + R_n^*(\boldsymbol{\theta}_j),$$

where $\mathbf{S}_j^* = \frac{1}{\sqrt{n\phi(h_0)}} \sum_{i:i \neq j} r_{ij} \mathbf{Z}_i^* K_{ij}$ and $R_n^*(\boldsymbol{\theta}_j) = \sum_{i:i \neq j} K_{ij} \int_0^{\eta_{ij}} \{I(\varepsilon_{i\tau} \leq \zeta_{ij} + t) - I(\varepsilon_{i\tau} \leq \zeta_{ij})\} dt$.

Consider the conditional expectation of $R_n^*(\boldsymbol{\theta}_j)$, we have

$$\mathbb{E}\{R_n^*(\boldsymbol{\theta}_j) | \mathbf{Z}, X\} = \sum_{i:i \neq j} K_{ij} \int_0^{\eta_{ij}} \{F_\tau(\zeta_{ij} + t | \mathbf{Z}, X) - F_\tau(\zeta_{ij} | \mathbf{Z}, X)\} dt = \frac{1}{2} \boldsymbol{\theta}_j^\top \mathbf{Q}_j \boldsymbol{\theta}_j + O_p(h_0^\alpha) + o_p(1).$$

where $\mathbf{Q}_j = \frac{1}{n\phi(h_0)} \sum_{i:i \neq j} f_\tau(0 | \mathbf{Z}, X) K_{ij} \mathbf{Z}_i^* \mathbf{Z}_i^{*\top}$. Using similar calculations, we can get $\text{Var}\{R_n^*(\boldsymbol{\theta}_j) | \mathbf{Z}, X\} = o_p(1)$. Therefore, we obtain

$$R_n^*(\boldsymbol{\theta}_j) = \mathbb{E}\{R_n^*(\boldsymbol{\theta}_j) | \mathbf{Z}, X\} + o_p(1) = \frac{1}{2} \boldsymbol{\theta}_j^\top \mathbf{Q}_j \boldsymbol{\theta}_j + o_p(1),$$

By Lemma A.1, we have

$$\mathbf{Q}_j = \mathbb{E}\{\mathbf{Q}_j\} + O_p\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\log n}{n}\right)}{n\phi(h_0)}}\right) = \mathbf{Q}_j^* + O_p\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\log n}{n}\right)}{n\phi(h_0)}}\right),$$

where $\mathbf{Q}_j^* = \frac{1}{\phi(h_0)} \mathbb{E}\left\{K(h_0^{-1}d(X, X_j)) \mathbb{E}\left[f_\tau(0 | \mathbf{Z}, X) \mathbf{Z}^* \mathbf{Z}^{*\top} | X\right]\right\}$. Thus, $\ell_n^*(\boldsymbol{\theta}_j)$ can be written as

$$\ell_n^*(\boldsymbol{\theta}_j) = \mathbf{S}_j^{*\top} \boldsymbol{\theta}_j + \frac{1}{2} \boldsymbol{\theta}_j^\top \mathbf{Q}_j^* \boldsymbol{\theta}_j + O_p\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\log n}{n}\right)}{n\phi(h_0)}} + h_0^\alpha\right).$$

Following from basic corollary of Hjort and Pollard (2011), the minimizer of $\ell_n^*(\boldsymbol{\theta}_j)$ can be written as

$$\tilde{\boldsymbol{\theta}}_j = -\{\mathbf{Q}_j^*\}^{-1} \mathbf{S}_j^* + O_p\left(\sqrt{\frac{\psi_{S_{\mathcal{F}}}\left(\frac{\log n}{n}\right)}{n\phi(h_0)}} + h_0^\alpha\right). \quad (8)$$

Proof of Theorem 1. Let $\boldsymbol{\theta} = \sqrt{n}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau)$, $\zeta_i = \tilde{a}_{\tau i} - m_\tau(X_i)$. By using the identity of Knight et al. (1998), we deduce that $\boldsymbol{\theta}$ is also the minimizer of function

$$\ell_n(\boldsymbol{\theta}) = \sum_{i=1}^n \left\{ \rho_\tau(\varepsilon_{i\tau} - \zeta_i - \mathbf{Z}_i^\top \boldsymbol{\theta} / \sqrt{n}) - \rho_\tau(\varepsilon_{i\tau} - \zeta_i) \right\} = \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n r_i \mathbf{Z}_i \right\}^\top \boldsymbol{\theta} + R_n(\boldsymbol{\theta}), \quad (9)$$

where $R_n(\boldsymbol{\theta}) = \sum_{i=1}^n \int_{\zeta_i}^{\zeta_i + \mathbf{Z}_i^\top \boldsymbol{\theta} / \sqrt{n}} \{I(\varepsilon_{i\tau} \leq t) - I(\varepsilon_{i\tau} \leq 0)\} dt$. Now we consider the conditional expectation of $R_n(\boldsymbol{\theta})$. It follows that

$$\mathbb{E}\{R_n(\boldsymbol{\theta})|\mathbf{Z}, X\} = \frac{1}{2}\boldsymbol{\theta}^\top \left(\sum_{i=1}^n \frac{1}{n} f_\tau(0|\mathbf{Z}, X) \mathbf{Z}_i \mathbf{Z}_i^\top \right) \boldsymbol{\theta} + \left(\frac{1}{\sqrt{n}} f_\tau(0|\mathbf{Z}, X) \sum_{i=1}^n \zeta_i \mathbf{Z}_i \right)^\top \boldsymbol{\theta} + o_p(1).$$

Using similar calculations, we have $\text{Var}\{R_n(\boldsymbol{\theta})|\mathbf{Z}, X\} = o_p(1)$. Thus, we deduce that

$$R_n(\boldsymbol{\theta}) = \mathbb{E}\{R_n(\boldsymbol{\theta})|\mathbf{Z}, X\} + o_p(1) = \frac{1}{2}\boldsymbol{\theta}^\top \mathbf{D}_n \boldsymbol{\theta} + \left(\frac{1}{\sqrt{n}} f_\tau(0|\mathbf{Z}, X) \sum_{i=1}^n \zeta_i \mathbf{Z}_i \right)^\top \boldsymbol{\theta} + o_p(1), \quad (10)$$

where $\mathbf{D}_n = \frac{1}{n} \sum_{i=1}^n f_\tau(0|\mathbf{Z}, X) \mathbf{Z}_i \mathbf{Z}_i^\top$. Combining (9) and (10), we have

$$\ell_n(\boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{\theta}^\top \mathbf{D}_n \boldsymbol{\theta} + \left(\frac{1}{\sqrt{n}} f_\tau(0|\mathbf{Z}, X) \sum_{i=1}^n \zeta_i \mathbf{Z}_i \right)^\top \boldsymbol{\theta} + \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n r_i \mathbf{Z}_i \right\}^\top \boldsymbol{\theta} + o_p(1),$$

By using (8), we can get

$$\begin{aligned} & \frac{1}{\sqrt{n}} f_\tau(0|\mathbf{Z}, X) \sum_{i=1}^n \zeta_i \mathbf{Z}_i \\ &= -\frac{1}{\sqrt{n}} f_\tau(0|\mathbf{Z}, X) \sum_{i=1}^n \mathbf{Z}_i (1, \mathbf{0}^\top) \frac{1}{\sqrt{n\phi(h_0)}} \left\{ \mathbb{E} \left\{ K(h_0^{-1}d(X, X_i)) \mathbb{E} \left[f_\tau(0|\mathbf{Z}, X) \mathbf{Z}^* \mathbf{Z}^{*\top} | X \right] \right\} \right\}^{-1} \\ & \quad \sqrt{\frac{\phi(h_0)}{n}} \sum_{j=1}^n \mathbf{Z}_j^* r_j K(h_0^{-1}d(X_i, X_j)) + O_p \left(\frac{1}{\phi(h_0)} \sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n}} + \frac{h_0^\alpha}{\sqrt{\phi(h_0)}} \right) \\ &= -\frac{1}{\sqrt{n}} \sum_{j=1}^n r_j \mathbf{A}_\tau(X_j, \mathbf{Z}_j) + o_p(1), \end{aligned}$$

Thus,

$$\ell_n(\boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{\theta}^\top \mathbf{D}_n \boldsymbol{\theta} + \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n r_i \{\mathbf{Z}_i - \mathbf{A}_\tau(X_i, \mathbf{Z}_i)\} \right\}^\top \boldsymbol{\theta} + o_p(1) = \frac{1}{2}\boldsymbol{\theta}^\top \mathbf{D}_n \boldsymbol{\theta} + \mathbf{S}_n^\top \boldsymbol{\theta} + o_p(1), \quad (11)$$

where $\mathbf{S}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n r_i \{\mathbf{Z}_i - \mathbf{A}_\tau(X_i, \mathbf{Z}_i)\}$. It is easy to show that $\mathbf{D}_n = \mathbb{E}\mathbf{D}_n + o_p(1) = \mathbf{M}_\tau + o_p(1)$. Substitute it to (11), we have $\ell_n(\boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{\theta}^\top \mathbf{M}_\tau \boldsymbol{\theta} + \mathbf{S}_n^\top \boldsymbol{\theta} + o_p(1)$. Following from basic corollary of Hjort and Pollard (2011), the minimizer of $\ell_n(\boldsymbol{\theta})$ can be written as $\hat{\boldsymbol{\theta}} = -\mathbf{M}_\tau^{-1} \mathbf{S}_n + o_p(1)$. By the Cramér-Wold Device, it is quite easy to show that the central limit theorem for \mathbf{S}_n holds. Furthermore, we have

$$\text{Var}(\mathbf{S}_n) = \tau(1-\tau) \mathbb{E}\{\mathbf{Z} - \mathbf{A}_\tau(X, \mathbf{Z})\} \{\mathbf{Z} - \mathbf{A}_\tau(X, \mathbf{Z})\}^\top = \tau(1-\tau) \mathbf{N}_\tau.$$

According to the central limit theorem, we have $\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau) \xrightarrow{\mathcal{L}} N(\mathbf{0}, \tau(1-\tau) \mathbf{M}_\tau^{-1} \mathbf{N}_\tau \mathbf{M}_\tau^{-1})$.

Proof of Theorem 2. By using the method similar to that used in Lemma A.2, we can show the conclusion similar to (8) holds, that is

$$\hat{\boldsymbol{\theta}} = -\hat{\mathbf{Q}}^{-1} \hat{\mathbf{S}}_1 + O_p \left(\sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)} + h^\alpha} \right)$$

uniformly for any $x \in S_{\mathcal{F}}$, where $\hat{\boldsymbol{\theta}} = \sqrt{n\phi(h)}(\hat{a}_\tau - m(x))$, $\zeta_i = m_\tau(x) - m_\tau(X_i)$ and $s_i = I\{\varepsilon_{i\tau} \leq \zeta_i + \mathbf{Z}_i^\top(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau)\} - \tau$, $\hat{\mathbf{S}}_1 = \frac{1}{\sqrt{n\phi(h)}} \sum_{i=1}^n K(h^{-1}d(X_i, x))s_i$. Thus,

$$\sup_{x \in S_{\mathcal{F}}} |\hat{a}_\tau - m(x)| = \sup_{x \in S_{\mathcal{F}}} |\hat{\mathbf{Q}}^{-1} \hat{\mathbf{S}}_2| + O_p \left(\frac{\sqrt{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}}{n\phi(h)} + \frac{h^\alpha}{\sqrt{n\phi(h)}} \right) \quad (12)$$

where $\hat{\mathbf{S}}_2 = \frac{1}{n\phi(h)} \sum_{i=1}^n K(h^{-1}d(X_i, x))s_i$, $\hat{\mathbf{Q}} = \frac{1}{\phi(h)} \mathbb{E} \{K(h^{-1}d(X, x))\mathbb{E}[f_\tau(0|\mathbf{Z}, X)|X]\}$. By using condition C9 and Fubini's Theorem, we have

$$\hat{\mathbf{Q}} = \frac{1}{\phi(h)} \mathbb{E} \{K(h^{-1}d(X, x))\mathbb{E}[f_\tau(0|\mathbf{Z}, X)|X]\} \asymp -\mathbb{E}[f_\tau(0|\mathbf{Z}, X)|X = x]. \quad (13)$$

Moreover, we have

$$\hat{\mathbf{S}}_2 = \frac{1}{n\phi(h)} \sum_{i=1}^n K(h^{-1}d(X_i, x))r_i + \frac{1}{n\phi(h)} \sum_{i=1}^n K(h^{-1}d(X_i, x))(s_i - r_i) = \hat{\mathbf{S}}_3 + \hat{\mathbf{S}}_4, \quad (14)$$

where $\hat{\mathbf{S}}_3 = \frac{1}{n\phi(h)} \sum_{i=1}^n K(h^{-1}d(X_i, x))r_i$, $\hat{\mathbf{S}}_4 = \frac{1}{n\phi(h)} \sum_{i=1}^n K(h^{-1}d(X_i, x))(s_i - r_i)$. Consider the conditional expectation of $\hat{\mathbf{S}}_4$, we have

$$\mathbb{E} \{ \hat{\mathbf{S}}_4 | \mathbf{Z}, X \} = \mathbb{E} \left\{ \frac{1}{n\phi(h)} \sum_{i=1}^n K(h^{-1}d(X_i, x))(s_i - r_i) | \mathbf{Z}, X \right\} = O_p(h^\alpha)$$

uniformly for any $x \in S_{\mathcal{F}}$. Using similar calculations, we have $\text{Var} \{ \hat{\mathbf{S}}_4 | \mathbf{Z}, X \} = o_p(h^{2\alpha})$. Thus, we have

$$\hat{\mathbf{S}}_4 = O_p(h^\alpha) \quad (15)$$

uniformly for any $x \in S_{\mathcal{F}}$. By Lemma A.1, we can obtain

$$\sup_{x \in S_{\mathcal{F}}} |\hat{\mathbf{S}}_3| = \mathbb{E} \hat{\mathbf{S}}_3 + O_p \left(\sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right) = O_p \left(\sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right) \quad (16)$$

According to condition C1, C3 and C9 and combining (12)-(16), we deduce that

$$\sup_{x \in S_{\mathcal{F}}} |\hat{a}_\tau - m_\tau(x)| = O_p \left(h^\alpha + \sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} + \frac{\sqrt{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}}{n\phi(h)} + \frac{h^\alpha}{\sqrt{n\phi(h)}} \right) = O_p \left(h^\alpha + \sqrt{\frac{\psi_{S_{\mathcal{F}}} \left(\frac{\log n}{n} \right)}{n\phi(h)}} \right).$$

References

- Aneiros, G., Ferraty, F. and Vieu, P. (2015). Variable selection in partial linear regression with functional covariate, *Statistics* **49**(6): 1322–1347.
- Aneiros-Pérez, G. and Vieu, P. (2006). Semi-functional partial linear regression, *Statistics & Probability Letters* **76**(11): 1102–1110.

- Aneiros-Pérez, G. and Vieu, P. (2008). Nonparametric time series prediction: A semi-functional partial linear modeling, *Journal of Multivariate Analysis* **99**(5): 834–857.
- Aneiros-Pérez, G. and Vieu, P. (2011). Automatic estimation procedure in partial linear model with functional data, *Statistical Papers* **52**(4): 751–771.
- Cardot, H., Crambes, C. and Sarda, P. (2005). Quantile regression when the covariates are functions, *Nonparametric Statistics* **17**(7): 841–856.
- Chen, K. and Müller, H.-G. (2012). Conditional quantile analysis when covariates are functions, with application to growth data, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **74**(1): 67–89.
- Cuevas, A. (2014). A partial overview of the theory of statistics with functional data, *Journal of Statistical Planning and Inference* **147**: 1–23.
- Ferraty, F., Laksaci, A., Tadj, A. and Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables, *Journal of statistical planning and inference* **140**(2): 335–352.
- Ferraty, F. and Vieu, P. (2006). *Nonparametric functional data analysis: theory and practice*, Springer Science & Business Media.
- Goia, A. and Vieu, P. (2014). Some advances in semiparametric functional data modelling, *Contributions in infinite-dimensional statistics and related topics. Esculapio, Bologna* pp. 135–141.
- Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics, *Journal of Multivariate Analysis* **146**: 1–6.
- Greven, S. and Scheipl, F. (2017). A general framework for functional regression modelling, *Statistical Modelling* **17**(1-2): 1–35.
- Hjort, N. L. and Pollard, D. (2011). Asymptotics for minimisers of convex processes, *arXiv preprint arXiv:1107.3806*.
- Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*, Vol. 200, Springer Science & Business Media.
- Hsing, T. and Eubank, R. L. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*, John Wiley & Sons.
- Kai, B., Li, R. and Zou, H. (2011). New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models, *Annals of statistics* **39**(1): 305.
- Kato, K. et al. (2012). Estimation in functional linear quantile regression, *The Annals of Statistics* **40**(6): 3108–3136.
- Knight, K. et al. (1998). Limiting distributions for L_1 regression estimators under general conditions, *The Annals of Statistics* **26**(2): 755–770.
- Kudraszow, N. L. and Vieu, P. (2013). Uniform consistency of knn regressors for functional variables, *Statistics & Probability Letters* **83**(8): 1863–1870.
- Ling, N., Aneiros, G. and Vieu, P. (2017). knn estimation in functional partial linear modeling, *Statistical Papers* pp. 1–22.
- Morris, J. S. (2015). Functional regression, *Annual Review of Statistics and Its Application* **2**(1): 321–359.
- Müller, H.-G. (2005). Functional modelling and classification of longitudinal data, *Scandinavian Journal of Statistics* **32**(2): 223–240.
- Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016). Functional data analysis, *Annual Review of Statistics and Its Application* **3**: 257–295.