Brazilian Journal of Probability and Statistics
2018, Vol. 32, No. 3, 616–656
https://doi.org/10.1214/17-BJPS357

Semiparametric quantile estimation for varying coefficient partially linear measurement errors models

Jun Zhang[a,b], Yan Zhou[a], Xia Cui[c] and Wangli Xu[d]

[a]*College of Mathematics and Statistics, Institute of Statistical Sciences, Shenzhen University*
[b]*Shenzhen-Hong Kong Joint Research Center for Applied Statistical Sciences, Shenzhen University*
[c]*School of Economics and Statistics, Guangzhou University*
[d]*Center for Applied Statistics, School of Statistics, Renmin University of China*

Abstract. We study varying coefficient partially linear models when some linear covariates are error-prone, but their ancillary variables are available. After calibrating the error-prone covariates, we study quantile regression estimates for parametric coefficients and nonparametric varying coefficient functions, and we develop a semiparametric composite quantile estimation procedure. Asymptotic properties of the proposed estimators are established, and the estimators achieve their best convergence rate with proper bandwidth conditions. Simulation studies are conducted to evaluate the performance of the proposed method, and a real data set is analyzed as an illustration.

1 Introduction

Various semiparametric regression models have become quite popular because they relax several restrictive assumptions on parametric models and remain flexible enough to capture the underlying relations between covariates and responses to dealing with real data. One of the most popular and important semiparametric regression models is the partial linear varying coefficient model (PLVCM):

$$Y = \boldsymbol{\gamma}_0^{\tau}\mathbf{Z} + \alpha_0(U) + \boldsymbol{\alpha}^{\tau}(U)\mathbf{X} + \varepsilon, \tag{1.1}$$

where "τ" denotes the transpose operator on a vector or a matrix throughout this paper. The variable Y is the response variable, and $\mathbf{Z}$, $\mathbf{X}$, and U are the covariates. $\boldsymbol{\gamma}_0$ is a vector of the unknown parameters, $\alpha_0(\cdot)$ and $\boldsymbol{\alpha}(\cdot)$ are unknown smooth functions, and ε is a random error with mean zero and finite variance. In this paper, we focus on univariate U only, although the proposed procedure is directly applicable for multivariate $\mathbf{U}$. However, the extension might be practically less useful due to the curse of dimensionality. Model (1.1) has attracted much attention due to the model's flexibility for combining multiple linear regression models and nonparametric regression models. The model includes important special cases. When $\boldsymbol{\alpha}(\cdot) \equiv 0$, we get partial linear models (PLMs), for example, Heckman (1986),

Key words and phrases. Ancillary variables, quantile regression, composite quantile regression, errors-in-variables, varying coefficient models.
Received December 2016; accepted February 2017.

Robinson (1988), Speckman (1988), Li, Feng and Peng (2011). Surveys of PLMs are given by Härdle, Liang and Gao (2000), Liang (2008). When $\boldsymbol{\gamma}_0 = 0$, model (1.1) reduces to varying coefficient models (VCMs), which have been applied to parsimoniously describe the data structure and to uncover scientific features, see Hastie and Tibshirani (1993), Cai, Fan and Li (2000), Fan and Zhang (2008), Wang and Xia (2009), Wei, Huang and Li (2011), Zhu, Li and Kong (2012), Gu and Liang (2014), Li, Ouyang and Racine (2013), Yuan et al. (2013).

Existing estimation procedures for the parameter $\boldsymbol{\gamma}_0$ and the nonparametric functions $\alpha_0(\cdot)$ and $\boldsymbol{\alpha}(\cdot)$ in PLVCMs are built on either least-squares or likelihood-based methods. For example, Zhang, Lee and Song (2002) proposed local polynomial estimators, which are further modified to achieve the optimal convergence rate. Xia, Zhang and Tong (2004) developed the idea of minimum average variance estimation (Xia et al., 2002, MAVE) and proposed an efficient estimation method. Li, Xue and Lian (2011) considered variable selection problem for PLVCMs when the number of parametric and nonparametric components diverge at appropriate rates. Bravo (2013) consider estimation and testing problems for PLVCMs when the response variable Y is missing at random.

It is known that ordinary least-squares regression is sensitive to extreme outliers, which distorts the results significantly. As a remedy, quantile regression (QR) is more robust against outliers in the response measurements, relative to ordinary least-squares regression. QR overcomes various problems that ordinary least-squares regression is usually confronted with and focuses on the relation between the response variable and covariates for a given quantile. QR has experienced deep and exciting developments in theory, methodology, and applications. Koenker (2005) provided a comprehensive survey of QR. For the estimation problem and statistical inference of semiparametric quantile regression (SQR), there is much work in the literature. See, for example, Hu, Gramacy and Lian (2013), Fan and Zhu (2013), Cai and Xiao (2012), Wang, Zhu and Zhou (2009). Recently, Kai, Li and Zou (2011) proposed an SQR procedure when the conditional quantile of the response-given covariates is modeled as PLVCMs, and they investigated the sampling properties of the proposed method. Kai, Li and Zou (2011) further proposed a semiparametric composite quantile regression (SCQR) for PLVCMs. The composite quantile regression (CQR) proposed by Zou and Yuan (2008) combines information across multiple quantile estimates to improve estimators, which are asymptotic efficient compared with the classical least-squares estimators. Kai, Li and Zou (2011) showed that the SCQR estimators for model (1.1) gain at least 88.9% efficiency for many non-normal errors and lose only a small amount of efficiency for normal errors. These properties have motivated researchers to develop CQR method for many other models; see Jiang et al. (2013), Guo et al. (2013), Jiang, Jiang and Song (2012), Kai, Li and Zou (2010). Recent discussions about the asymptotic relative efficiency are referred to by Feng, Zou and Wang (2012), Kai, Li and Zou (2010), Sun, Gai and Lin (2013), Shang, Zou and Wang (2012), Wang, Kai and Li (2009).

In many biomedical studies, covariates may be observed with certain contamination. As we know, measurement errors in covariates may cause a large bias for the regression coefficients if we ignore measurement errors. Fuller (1987) and Carroll et al. (2006) are two comprehensive surveys of linear and nonlinear measurement errors models. In the literature of errors-in-variables for QR, He and Liang (2000) considered the QR procedure for partial linear errors-in-variables models. Wang, Stefanski and Zhu (2012) developed a corrected score to account for a class of covariates measurement errors in QR. This method is simple to implement and does not need parametric assumptions of the regression errors. Wei and Carroll (2009) constructed joint estimating equations that simultaneously hold for all the quantile levels, which produces a consistent linear quantile estimator by correcting bias caused by the measurement errors. In this paper, we consider a scenario where p $(1 \le p \le d)$ components of $\mathbf{Z}$, namely, $\boldsymbol{\xi}$, are observed with ancillary variables $\boldsymbol{\eta}$ and V through the following model:

$$\boldsymbol{\eta} = \boldsymbol{\xi}(V) + \mathbf{e}, \tag{1.2}$$

where $\mathbf{e}$ is the model error with $E(\mathbf{e}|V) = 0$ and finite covariance matrix $\boldsymbol{\Sigma}_{\mathbf{e}} = \mathrm{Cov}(\mathbf{e})$. We focus on univariate V only, although the proposed procedure in this paper is directly applicable for multivariate $\mathbf{V}$. However, the extension might be practically less useful due to the curse of dimensionality. This kind of measurement errors model is not uncommon and includes various models, for example, the de-noise linear or nonlinear models studied by Cui, He and Zhu (2002), Cui and Hu (2011), Cai, Naik and Tsai (2000), the rational expectation model in the econometric literature, and the errors-in-variables model for the Duchenne muscular dystrophy (DMD) study considered in Zhou and Liang (2009), who studied the estimation and hypothesis testing problems for the models (1.1)–(1.2) by using the least-squares estimation method.

The QR provides a more complete picture of the conditional distribution of responses given the covariates when the lower and upper or all quantiles are of interest and has an advantage over ordinary least-squares regression due to its flexibility for modeling data with heterogeneous conditional distributions. This motivates us to study the QR estimation in the case of measurement errors model (1.2). In Section 2, we propose a semiparametric quantile measurement errors regression. We investigate the sampling properties of the proposed methods and their asymptotic normality. Under proper conditions for the bandwidths, we show that the estimators of the parameters are root-n consistent, and the quantile regression estimators for the nonparametric parts achieve the optimal rate of convergence. In Section 3, we also propose SCQR estimators to estimate the parameters and nonparametric components in PLVCMs in the case of measurement errors model (1.2). The asymptotic properties of the SCQR estimators are also presented. Simulation results and a real data analysis are presented in Section 4. The regularity conditions and technical proofs are given in the Appendix.

2 Semiparametric quantile regression

Let Y be the response variable, $\mathbf{X} = (X_1, \dots, X_p)^\tau \in \mathbb{R}^p$, $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)^\tau \in \mathbb{R}^d$, $\mathbf{W} = (W_1, \dots, W_r)^\tau \in \mathbb{R}^r$, and $U \in \mathbb{R}^1$ be the covariates. The semiparametric quantile varying coefficient partial linear model (SQVCPLM) assumes that the conditional quantile function of the Y-given covariates $(\boldsymbol{\xi}, \mathbf{W}, \mathbf{X}, U)$, $Q_\kappa(\xi, \mathbf{w}, \mathbf{x}, u) = \arg\min_a E\{\rho_\kappa(Y - a) | \boldsymbol{\xi} = \xi, \mathbf{W} = \mathbf{w}, \mathbf{X} = \mathbf{x}, U = u\}$ is expressed as

$$Q_\kappa(\xi, \mathbf{w}, \mathbf{x}, u) = \boldsymbol{\beta}_\kappa^\tau \xi + \boldsymbol{\theta}_\kappa^\tau \mathbf{w} + \alpha_{0,\kappa}(u) + \boldsymbol{\alpha}_\kappa(u)^\tau \mathbf{x}, \tag{2.1}$$

where $\rho_\kappa(t) = \kappa t - t I\{t < 0\}$ is the quantile loss function at $\kappa \in (0, 1)$.

Here, it is assumed that the components of $\boldsymbol{\xi}$ are unobserved, but auxiliary variables $(\boldsymbol{\eta}, V)$ are available to remit $\boldsymbol{\xi}$. Moreover, the observed variable $\boldsymbol{\eta}$ is related to the observed variable V via

$$\boldsymbol{\eta} = \boldsymbol{\xi}(V) + \mathbf{e}, \tag{2.2}$$

where $\mathbf{e}$ is a measurement error and independent of $(\mathbf{X}, \mathbf{W}, V, U, Y)$. We call model (2.1) and model (2.2) as semiparametric quantile varying coefficient partial linear measurement errors models (SQVCPLMeMs).

2.1 Covariate calibration

Suppose that $\{Y_i, U_i, \boldsymbol{\eta}_i, V_i, \mathbf{W}_i, \mathbf{X}_i, i = 1, \dots, n\}$ be an independent and identically distributed sample from $(Y, U, \boldsymbol{\eta}, V, \mathbf{W}, \mathbf{X})$. When the covariate $\boldsymbol{\xi}$ is measured with errors, we first calibrate $\boldsymbol{\xi}$ by using the ancillary observed sample $\{\boldsymbol{\eta}_i, V_i, i = 1, \dots, n\}$.

Now, we introduce the calibration estimation procedure to remit $\boldsymbol{\xi}$. Let $\boldsymbol{\eta}_{i,k}$ be the kth entry of vector $\boldsymbol{\eta}_i$ for $i = 1, \dots, n$. The local linear smoothing technique (Fan and Gijbels, 1996) is applied to estimate $\xi_k(v)$, the kth component of $\boldsymbol{\xi}(v)$. That is, to minimize

$$\sum_{i=1}^n \{\boldsymbol{\eta}_{i,k} - a_{0k} - a_{1k}(V_i - v)\}^2 L_{h_k}(V_i - v), \tag{2.3}$$

with respect to a_{0k}, a_{1k}, where $L_{h_k}(\cdot) = L(\cdot/h_k)/h_k$ with $L(\cdot)$ is a kernel function, h_k $(k = 1, \dots, d)$ is the bandwidth. Let $\hat{a}_{0k}$ and $\hat{a}_{1k}$ be the minimizers of (2.3), we have

$$\hat{\xi}_k(v) = \hat{a}_{0k} = \frac{A_{20,k}(v) A_{01,k}(v) - A_{10,k}(v) A_{11,k}(v)}{A_{00,k}(v) A_{20,k}(v) - A_{10,k}^2(v)}, \tag{2.4}$$

where $A_{j_1 j_2,k}(v) = \sum_{i=1}^n L_{h_k}(V_i - v)(V_i - v)^{j_1} \boldsymbol{\eta}_{i,k}^{j_2}$ for $j_1 = 0, 1, 2$, $j_2 = 0, 1$, $k = 1, \dots, d$. Under the conditions given in the Appendix (see also in Zhou and Liang

(2009)), we have the following asymptotic expression:

$$\begin{aligned}\hat{\xi}_k(v) - \xi_k(v) \\ = \frac{\mu_{L_2}}{2} h_k^2 \xi_k^{(2)}(v) + \frac{1}{n f_{\mathrm{V}}(v)} \sum_{i=1}^n L_{h_k}(V_i - v)\mathbf{e}_{i,k} + o\big(h_k^2 + \log h_k^{-1}/\sqrt{n h_k}\big),\end{aligned} \tag{2.5}$$

uniformly on $v \in \mathbb{V}$, where $\mathbb{V}$ is a bounded support of V. Here $\xi_k^{(2)}(v)$ is the second derivative of $\xi_k(v)$, $\mathbf{e}_{i,k}$ is the kth component of $\mathbf{e}_i$, $i = 1, \dots, n$. $\mu_{L_2} = \int u^2 L(u)\, du$, and $f_V(v)$ is the density of V.

2.2 Estimation and main results

After calibrating $\boldsymbol{\xi}$, we model the synthesis data $\{Y_i, U_i, \hat{\boldsymbol{\xi}}_i, \mathbf{W}_i, \mathbf{X}_i; i = 1, \dots, n\}$ by using the SQR principle (Kai, Li and Zou, 2011):

$$Y \approx \boldsymbol{\beta}_\kappa^\tau \hat{\boldsymbol{\xi}} + \boldsymbol{\theta}_\kappa^\tau \mathbf{W} + \alpha_{0,\kappa}(U) + \boldsymbol{\alpha}_\kappa(U)^\tau \mathbf{X} + \varepsilon_\kappa, \tag{2.6}$$

where ε_κ is the random error with the conditional κth quantile zero. The quantile estimators of $\boldsymbol{\beta}_\kappa, \boldsymbol{\theta}_\kappa, \alpha_{0,\kappa}(\cdot)$, and $\boldsymbol{\alpha}_\kappa(\cdot)$ are obtained by minimizing the quantile loss function:

$$\sum_{i=1}^n \rho_\kappa\big\{Y_i - \alpha_{0,\kappa}(U_i) - \boldsymbol{\alpha}_\kappa^\tau(U_i)\mathbf{X}_i - \boldsymbol{\theta}_\kappa^\tau \mathbf{W}_i - \boldsymbol{\beta}_\kappa^\tau \hat{\boldsymbol{\xi}}_i\big\}. \tag{2.7}$$

As Kai, Li and Zou (2011), Li and Liang (2008) claimed, different convergence rates of the parametric components $\boldsymbol{\beta}_\kappa$ and $\boldsymbol{\theta}_\kappa$ and the nonparametric components $\alpha_{0,\kappa}(\cdot)$ and $\boldsymbol{\alpha}_\kappa(\cdot)$ are involved in (2.7). A three-stage estimation procedure is proposed to obtain the proper estimators. We first use a local linear smoothing technique (Fan and Gijbels, 1996) to approximate $\alpha_{0,\kappa}(\cdot)$ and $\boldsymbol{\alpha}_\kappa(\cdot)$ and obtain the initial local minimizer of $(\boldsymbol{\beta}_\kappa^\tau, \boldsymbol{\theta}_\kappa^\tau, \alpha_{0,\kappa}(\cdot), \boldsymbol{\alpha}_\kappa^\tau(\cdot))^\tau$. In the second and third stages, we use these initial estimators alternatively to obtain refined estimators of $(\boldsymbol{\beta}_\kappa^\tau, \boldsymbol{\theta}_\kappa^\tau)^\tau$ and $(\alpha_{0,\kappa}(\cdot), \boldsymbol{\alpha}_\kappa^\tau(\cdot))^\tau$.

Let $K(\cdot)$ be the kernel function, h be the bandwidth, and $K_h(\cdot) = K(\cdot/h)/h$. Recall that $\boldsymbol{\alpha}_\kappa(u) = (\alpha_{1,\kappa}(u), \dots, \alpha_{p,\kappa}(u))^\tau$, $\boldsymbol{\alpha}'(u) = (\alpha'_{1,\kappa}(u), \dots, \alpha'_{p,\kappa}(u))^\tau$. For each u in a neighborhood of U, we approximate $\alpha_{j,\kappa}(U)$ by $\alpha_{j,\kappa}(u) + \alpha'_{j,\kappa}(u)(U - u)$, $j = 0, 1, \dots, p$. The local estimators of $\boldsymbol{\beta}_\kappa$, $\boldsymbol{\theta}_\kappa$, $\alpha_{j,\kappa}(u)$, and $\alpha'_{j,\kappa}(u)$ are obtained by minimizing the following local quantile loss function with respect to α_0, α'_0, $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}'$, $\boldsymbol{\beta}$, and $\boldsymbol{\theta}$:

$$\begin{aligned}\sum_{i=1}^n \rho_\kappa\big\{Y_i - \boldsymbol{\beta}^\tau \hat{\boldsymbol{\xi}}_i - \boldsymbol{\theta}^\tau \mathbf{W}_i - \alpha_0 - \alpha'_0(U_i - u) - \boldsymbol{\alpha}^\tau \mathbf{X}_i - \boldsymbol{\alpha}'^\tau \mathbf{X}_i(U_i - u)\big\} \\ \times K_h(U_i - u).\end{aligned} \tag{2.8}$$

Denote the local estimators of $\alpha_0, \boldsymbol{\alpha}, \boldsymbol{\beta}$, and $\boldsymbol{\theta}$ from (2.8) by $\breve{\alpha}_{0,\kappa}(u), \breve{\boldsymbol{\alpha}}_\kappa(u), \breve{\boldsymbol{\beta}}_\kappa$, and $\breve{\boldsymbol{\theta}}_\kappa$. As demonstrated in Theorem 1 in the following, these estimators are all $\sqrt{nh}$-consistent.

Denote $\mathrm{A}^{\otimes 2} = \mathrm{A}\mathrm{A}^\tau$ for any matrix or vector A. Let $f_\kappa(\cdot|v, w, x, u)$ be the density function of the error ε_κ conditional on $(V, \mathbf{W}, \mathbf{X}, U) = (v, \mathbf{w}, \mathbf{x}, u)$ and $f_U(\cdot)$ be the marginal density function of the covariate U. Define $\mathbf{M} = (\boldsymbol{\xi}^\tau, \mathbf{W}^\tau, 1, \mathbf{X}^\tau)^\tau$ and

$$\mathbf{A}_1(u) = E\big[\mathbf{M}^{\otimes 2} f_\kappa(0|V, \mathbf{W}, \mathbf{X}, U)|U = u\big], \qquad \mathbf{B}_1(u) = E\big[\mathbf{M}^{\otimes 2}|U = u\big],$$
$$\boldsymbol{\Upsilon}(u) = E\big[\mathbf{M}(\boldsymbol{\xi}^{(2)}(V))^\tau f_\kappa(0|V, \mathbf{W}, \mathbf{X}, U)|U = u\big]\mathbf{C}\boldsymbol{\beta}_\kappa,$$

where $\boldsymbol{\xi}^{(2)}(V) = (\xi_1^{(2)}(V), \ldots, \xi_d^{(2)}(V))^\tau$, $\mathbf{C} = \mathrm{diag}(c_1^2, \ldots, c_d^2)$, and c_i are defined in condition (C5) in the Appendix. Moreover,

$$\mu_{K_2} = \int t^2 K(t)\,dt, \qquad \mu_{L_2} = \int t^2 L(t)\,dt, \qquad \vartheta_{K_0} = \int K^2(t)\,dt.$$

Theorem 1. *Under the regularity conditions* (C1)–(C4), (C5)(i), *and* (C6) *given in the Appendix, we have*

$$\sqrt{nh}\left\{\begin{pmatrix} \breve{\boldsymbol{\beta}}_\kappa - \boldsymbol{\beta}_\kappa \\ \breve{\boldsymbol{\theta}}_\kappa - \boldsymbol{\theta}_\kappa \\ \breve{\alpha}_{0,\kappa}(u) - \alpha_{0,\kappa}(u) \\ \breve{\boldsymbol{\alpha}}_\kappa(u) - \boldsymbol{\alpha}_\kappa(u) \end{pmatrix} - \frac{h^2\mu_{K_2}}{2}\begin{pmatrix} \mathbf{0}_{d\times 1} \\ \mathbf{0}_{r\times 1} \\ \alpha''_{0,\kappa}(u) \\ \boldsymbol{\alpha}''_\kappa(u) \end{pmatrix} + \frac{h_o^2\mu_{L_2}}{2}\mathbf{A}_1^{-1}(u)\boldsymbol{\Upsilon}(u)\right\}$$
$$\xrightarrow{\mathcal{L}} N\left(\mathbf{0}, \frac{\kappa(1-\kappa)\vartheta_{K_0}}{f_U(u)}\mathbf{A}_1^{-1}(u)\mathbf{B}_1(u)\mathbf{A}_1^{-1}(u)\right).$$

Remark 1. As the data is used in a neighborhood of u based on the local loss function (2.8), Theorem 1 indicates that the estimators $\breve{\boldsymbol{\beta}}_\kappa$ and $\breve{\boldsymbol{\theta}}_\kappa$ are all $\sqrt{nh}$-consistent with an extra bias $\boldsymbol{\pi}_1\frac{h_o^2\mu_{L_2}}{2}\mathbf{A}_1^{-1}(u)\boldsymbol{\Upsilon}(u)$, $\boldsymbol{\pi}_1 = (\mathbf{I}_{d+r}, \mathbf{0})_{(d+r)\times(d+r+1+p)}$, which is caused due to the estimation of unobserved $\boldsymbol{\xi}(v)$. This extra bias will converge to zero as $nhh_o^4 \to 0$. If $\boldsymbol{\xi}$ is observed exactly, then the bias term $\frac{h_o^2\mu_{L_2}}{2}\mathbf{A}_1^{-1}(u)\boldsymbol{\Upsilon}(u)$ vanishes, and the asymptotic result in Theorem 1 is the same as Theorem 2.1 obtained in Kai, Li and Zou (2011).

After the first-stage estimators $\breve{\alpha}_{0,\kappa}(u)$, and $\breve{\boldsymbol{\alpha}}_\kappa(u)$ are obtained, the improved estimators for $\boldsymbol{\beta}_\kappa, \boldsymbol{\theta}_\kappa$ are obtained by minimizing the global objective function with respect to $(\boldsymbol{\beta}, \boldsymbol{\theta})$:

$$\sum_{i=1}^n \rho_\kappa\big\{Y_i - \boldsymbol{\beta}^\tau\hat{\boldsymbol{\xi}}_i - \boldsymbol{\theta}^\tau\mathbf{W}_i - \breve{\alpha}_{0,\kappa}(U_i) - \breve{\boldsymbol{\alpha}}^\tau(U_i)\mathbf{X}_i\big\}. \tag{2.9}$$

Denote the global estimators from (2.9) by $\hat{\boldsymbol{\beta}}_\kappa$ and $\hat{\boldsymbol{\theta}}_\kappa$. Theorem 3 presents the root-n asymptotic normality of $\hat{\boldsymbol{\beta}}_\kappa$ and $\hat{\boldsymbol{\theta}}_\kappa$. Define

$$\boldsymbol{\lambda}_\kappa(v) = E\big[(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau)^\tau f_\kappa(0|V, \mathbf{W}, \mathbf{X}, U)|V = v\big],$$

$$\boldsymbol{\phi}_\kappa(u) = E\big[(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau)^\tau (\mathbf{0}^\tau_{(d+r)\times 1}, 1, \mathbf{X}^\tau) f_\kappa(0|V, \mathbf{W}, \mathbf{X}, U)|U = u\big]\mathbf{A}_1^{-1}(u).$$

We have the following asymptotic result.

Theorem 2. *Under the regularity conditions* (C1)–(C4), (C5)(ii), *and* (C6) *given in the Appendix, we have*

$$\sqrt{n}\begin{pmatrix}\hat{\boldsymbol{\beta}}_\kappa - \boldsymbol{\beta}_\kappa \\ \hat{\boldsymbol{\theta}}_\kappa - \boldsymbol{\theta}_\kappa\end{pmatrix} \xrightarrow{\mathcal{L}} N\big(\mathbf{0}_{(d+r)\times 1}, \boldsymbol{\Lambda}_\kappa^{-1}\boldsymbol{\Sigma}_\kappa\boldsymbol{\Lambda}_\kappa^{-1}\big),$$

where

$$\boldsymbol{\Lambda}_\kappa = E\big[\{(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau)^\tau\}^{\otimes 2} f_\kappa(0|V, \mathbf{W}, \mathbf{X}, U)\big],$$

$$\boldsymbol{\Sigma}_\kappa = E\big[\boldsymbol{\lambda}_\kappa^{\otimes 2}(V)\big]\boldsymbol{\beta}_\kappa^\tau\boldsymbol{\Sigma}_{\mathbf{e}}\boldsymbol{\beta}_\kappa + E\big[\{(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau)^\tau + \boldsymbol{\phi}_\kappa(U)\mathbf{M}\}^{\otimes 2}\big]\kappa(1-\kappa).$$

Remark 2. The proposed estimation procedure (2.9) involves bandwidths, h and h_k, $k = 1, \ldots, d$, to be selected. By condition (C5)(ii) of that $nh^4 \to 0$ and $nh_k^4 \to 0$, under-smoothing is necessary for obtaining root-n consistent estimators of $\boldsymbol{\beta}_\kappa, \boldsymbol{\theta}_\kappa$. Carroll et al. (1997) suggested that the rate for the undersmoothing bandwidth is an order of $O(n^{-1/5}) \times n^{-2/15} = O(n^{-1/3})$, and this rate meets the requirements of condition (C5)(ii). In practice, those useful and reasonable choices for h and h_k are implemented by Zhou and Liang (2009); i.e., $h = \hat{\sigma}_U n^{-1/3}$, $h_k = \hat{\sigma}_V n^{-1/3}$, where $\hat{\sigma}_U$ and $\hat{\sigma}_V$ are the sample deviations of U and V, respectively.

The first term $E[\boldsymbol{\lambda}_\kappa^{\otimes 2}(V)]\boldsymbol{\beta}_\kappa^\tau\boldsymbol{\Sigma}_{\mathbf{e}}\boldsymbol{\beta}_\kappa$ in the asymptotic variance $\boldsymbol{\Sigma}_\kappa$ is caused by estimating unobserved $\boldsymbol{\xi}(v)$. The rest of $\boldsymbol{\Sigma}_\kappa$ is the same as Theorem 2.2 in Kai, Li and Zou (2011) if $\boldsymbol{\xi}(v) = v$ and measurement error $\mathbf{e}$ vanishes (i.e., $\mathbf{e} = \mathbf{0}$).

Next, we improve the estimation efficiency of $\alpha_{0,\kappa}(u)$ and $\boldsymbol{\alpha}_\kappa(u)$ by using the root-n consistent estimators $\hat{\boldsymbol{\beta}}_\kappa$ and $\hat{\boldsymbol{\theta}}_\kappa$. Let $\hat{\alpha}_{0,\kappa}(u), \widehat{\alpha'}_{0,\kappa}(u), \hat{\boldsymbol{\alpha}}_\kappa(u),$ and $\widehat{\boldsymbol{\alpha}'}_\kappa(u)$ be the minimizers of

$$\sum_{i=1}^n \rho_\kappa\{Y_i - \hat{\boldsymbol{\beta}}^\tau\hat{\boldsymbol{\xi}}_i - \hat{\boldsymbol{\theta}}^\tau\mathbf{W}_i - \alpha_0 - b_0(U_i - u) - \boldsymbol{\alpha}^\tau\mathbf{X}_i - \mathbf{b}^\tau\mathbf{X}_i(U_i - u)\} \times K_h(U_i - u), \tag{2.10}$$

with respect to $\alpha_0, b_0, \boldsymbol{\alpha},$ and $\mathbf{b}$. We have the following asymptotic result.

Theorem 3. *Under the regularity conditions* (C1)–(C4), (C5)(i), *and* (C6) *given in the Appendix, we have*

$$\sqrt{nh}\left\{\begin{pmatrix}\hat{\alpha}_{0,\kappa}(u)-\alpha_{0,\kappa}(u)\\ \hat{\boldsymbol{\alpha}}_{\kappa}(u)-\boldsymbol{\alpha}_{\kappa}(u)\end{pmatrix}-\frac{h^2\mu_{K_2}}{2}\begin{pmatrix}\alpha''_{0,\kappa}(u)\\ \boldsymbol{\alpha}''_{\kappa}(u)\end{pmatrix}+\frac{h_o^2\mu_{L_2}}{2}\mathbf{A}_2^{-1}(u)\boldsymbol{\Psi}(u)\right\}$$
$$\xrightarrow{\mathcal{L}} N\left(\mathbf{0}_{(p+1)\times 1},\frac{\kappa(1-\kappa)\vartheta_{K_0}}{f_U(u)}\mathbf{A}_2^{-1}(u)\mathbf{B}_2(u)\mathbf{A}_2^{-1}(u)\right),$$

where

$$\mathbf{A}_2(u)=E\big[\big[(1,\mathbf{X}^\tau)^\tau\big]^{\otimes 2}f_\kappa(0|V,\mathbf{W},\mathbf{X},U)|U=u\big],$$
$$\mathbf{B}_2(u)=E\big[\big[(1,\mathbf{X}^\tau)^\tau\big]^{\otimes 2}|U=u\big],$$
$$\boldsymbol{\Psi}(u)=E\big[(1,\mathbf{X}^\tau)^\tau(\boldsymbol{\xi}^{(2)}(V))^\tau f_\kappa(0|V,\mathbf{W},\mathbf{X},U)|U=u\big]\mathbf{C}\boldsymbol{\beta}_\kappa.$$

Remark 3. The extra bias $\frac{h_o^2\mu_{L_2}}{2}\mathbf{A}_2^{-1}(u)\boldsymbol{\Psi}(u)$ is due to the estimation of unobserved $\boldsymbol{\xi}$. If $\boldsymbol{\xi}$ is observed exactly, then this bias term will vanish, and the asymptotic result in Theorem 3 is the same as Theorem 2.3 obtained in Kai, Li and Zou (2011). As noted in Kai, Li and Zou (2011), the refined estimators for $\alpha_{0,\kappa}(u)$ and $\boldsymbol{\alpha}_\kappa(u)$ obtained by (2.10) have smaller conditional asymptotic variances than $\breve{\alpha}_{0,\kappa}(u)$ and $\breve{\boldsymbol{\alpha}}_\kappa(u)$ obtained by (2.8). In this context, the refined estimation procedure (2.10) provides more efficient estimators for these unknown quantities $\alpha_{0,\kappa}(u)$ and $\boldsymbol{\alpha}_\kappa(u)$.

3 Semiparametric composite quantile estimation

In this section, we aim to develop an SCQR estimate under the following measurement error setting:

$$\begin{cases}Y=\boldsymbol{\beta}^\tau\boldsymbol{\xi}+\boldsymbol{\theta}^\tau\mathbf{W}+\alpha_0(U)+\boldsymbol{\alpha}(U)^\tau\mathbf{X}+\varepsilon,\\ \boldsymbol{\eta}=\boldsymbol{\xi}(V)+\mathbf{e}.\end{cases}\tag{3.1}$$

The random error ε is independent with $(\boldsymbol{\xi}^\tau,\mathbf{W}^\tau,U,\mathbf{X}^\tau)^\tau$ and independent with $(V,\mathbf{e}^\tau)^\tau$. We assume that ε follows the distribution $F_\varepsilon(\cdot)$ with mean zero. If $\boldsymbol{\xi}$ is observed, for a given $\kappa_l\in(0,1)$, we have

$$Q_\kappa(\xi,\mathbf{w},\mathbf{x},u)=c_{\kappa_l}+\boldsymbol{\beta}^\tau\xi+\boldsymbol{\theta}^\tau\mathbf{w}+\alpha_0(u)+\boldsymbol{\alpha}(u)^\tau\mathbf{x},$$

where $c_{\kappa_l}=F_\varepsilon^{-1}(\kappa_l)$. As Kai, Li and Zou (2011), Zou and Yuan (2008), Kai, Li and Zou (2010) indicated, the local CQR method significantly improves the estimation efficiency compared with the local least-squares estimator when the model error ε follows non-normal distributions. This motivates us to investigate the SCQR estimate for the SQVCPLMeMs introduced in model (3.1).

Using those calibrated variables $\{Y_i, \hat{\boldsymbol{\xi}}_i, \mathbf{W}_i, \mathbf{X}_i, U_i, i = 1, \ldots, n\}$, the SCQR estimators of $\boldsymbol{\beta}$, $\boldsymbol{\theta}$, $\alpha_0(u)$, and $\boldsymbol{\alpha}(u)$ in model (3.1) are obtained by minimizing the following objective function (3.2):

$$\sum_{l=1}^{q}\sum_{i=1}^{n}\rho_{\kappa_l}\big(Y_i - \alpha_{0l}(U_i) - \boldsymbol{\alpha}(U_i)^\tau \mathbf{X}_i - \boldsymbol{\beta}^\tau \hat{\boldsymbol{\xi}}_i - \boldsymbol{\theta}^\tau \mathbf{W}_i\big), \tag{3.2}$$

where $\kappa_l = \frac{l}{q+1}$ for a given $q, l = 1, \ldots, q$. We now adapt the three-stage estimation procedure proposed in Section 2.1.

Let $\boldsymbol{\alpha}(u) = (\alpha_1(u), \ldots, \alpha_p(u))^\tau$, $\boldsymbol{\alpha}'(u) = (\alpha_1'(u), \ldots, \alpha_p'(u))^\tau$. For each u in a neighborhood of U, we approximate $\alpha_j(U)$ by $\alpha_j(u) + \alpha_j'(u)(U - u)$, $j = 0, 1, \ldots, p$. The local estimates of $\boldsymbol{\beta}$, $\boldsymbol{\theta}$, $\alpha_j(u)$, and $\alpha_j'(u)$ are obtained by minimizing the following local composite quantile loss function with respect to α_{0l}^*, $\alpha_0'^*$, $\boldsymbol{\alpha}^*$, $\boldsymbol{\alpha}'^*$, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, $l = 1, \ldots, q$,

$$\begin{aligned}\sum_{l=1}^{q}\sum_{i=1}^{n}\rho_{\kappa_l}\big\{&Y_i - \boldsymbol{\beta}^{*\tau}\hat{\boldsymbol{\xi}}_i - \boldsymbol{\theta}^{*\tau}\mathbf{W}_i - \alpha_{0l}^* - \alpha_0'^*(U_i - u)\\ &- \boldsymbol{\alpha}^{*\tau}\mathbf{X}_i - \boldsymbol{\alpha}'^{*\tau}\mathbf{X}_i(U_i - u)\big\}\\ &\times K_h(U_i - u).\end{aligned} \tag{3.3}$$

Denote the local estimators of $\alpha_{0l}^*, \boldsymbol{\alpha}^*, \boldsymbol{\beta}$, and $\boldsymbol{\theta}$ from (3.3) by $\breve{\alpha}_{0l}(u), \breve{\boldsymbol{\alpha}}(u), \breve{\boldsymbol{\beta}}, \breve{\boldsymbol{\theta}}$, $l = 1, \ldots, q$. The estimator of $\alpha_0(u)$ is given by

$$\breve{\alpha}_0(u) = \frac{1}{q}\sum_{l=1}^{q}\breve{\alpha}_{0l}(u). \tag{3.4}$$

In Theorem 4, we present that these estimators are all $\sqrt{nh}$-consistent.

Theorem 4. *Under the regularity conditions* (C1)–(C4), (C5)(i), *and* (C7) *given in the Appendix, we have*

$$\sqrt{nh}\left\{\begin{pmatrix}\breve{\boldsymbol{\beta}} - \boldsymbol{\beta}\\ \breve{\boldsymbol{\theta}} - \boldsymbol{\theta}\\ \breve{\alpha}_{01}(u) - \alpha_0(u) - c_{\kappa_1}\\ \vdots\\ \breve{\alpha}_{0q}(u) - \alpha_0(u) - c_{\kappa_q}\\ \breve{\boldsymbol{\alpha}}(u) - \boldsymbol{\alpha}(u)\end{pmatrix} - \frac{h^2\mu_{K_2}}{2}\begin{pmatrix}\mathbf{0}_{d\times 1}\\ \mathbf{0}_{r\times 1}\\ \mathbf{1}_{q\times 1}\alpha_0''(u)\\ \boldsymbol{\alpha}''(u)\end{pmatrix} + \frac{h_o^2\mu_{L_2}}{2}\mathbf{A}_{1,f_\varepsilon}^{-1}(u)\aleph(u)\right\}$$

$$\xrightarrow{\mathcal{L}} N\left(\mathbf{0}_{(d+r+q+p)\times 1}, \frac{\vartheta_{K_0}}{f_U(u)}\mathbf{A}_{1,f_\varepsilon}^{-1}(u)\mathbf{T}(u)\mathbf{A}_{1,f_\varepsilon}^{-1}(u)\right),$$

where

$$\mathbf{A}_{1,f_\varepsilon}(u) = \sum_{l=1}^{q} f_\varepsilon(c_{\kappa_l})E\big[\mathbf{M}_{[l]}^{\otimes 2}|U = u\big],$$

$$\aleph(u) = \sum_{l=1}^{q} f_\varepsilon(c_{\kappa_l}) E\big[\mathbf{M}_{[l]}\big(\boldsymbol{\xi}^{(2)}(V)\big)^\tau | U = u\big]\mathbf{C}\boldsymbol{\beta},$$

$$\mathbf{T}(u) = \sum_{l=1}^{q}\sum_{s=1}^{q} E\big[\mathbf{M}_{[l]}\mathbf{M}_{[s]}^\tau | U = u\big](\kappa_l \wedge \kappa_s - \kappa_l\kappa_s),$$

and $\mathbf{M}_{[l]} = (\boldsymbol{\xi}^\tau, \mathbf{W}^\tau, m_l^\tau, \mathbf{X}^\tau)^\tau$. *Here,* m_l *is a* q*-vector with* 1 *at the* l*th position and* 0 *elsewhere.*

Remark 4. Theorem 4 reveals that the local SCQR estimation procedure (3.3) also entails a $\sqrt{nh}$-consistent estimator of $(\boldsymbol{\beta}^\tau, \boldsymbol{\theta}^\tau)^\tau$ with an extra bias $\boldsymbol{\pi}^* \frac{h_o^2 \mu_{L_2}}{2} \times \mathbf{A}_{1,f_\varepsilon}^{-1}(u)\aleph(u)$, $\boldsymbol{\pi}^* = (\mathbf{I}_{d+r}, \mathbf{0})_{(d+r)\times(d+r+q+p)}$. Analogous with Theorem 1, this bias $\frac{h_o^2\mu_{L_2}}{2}\mathbf{A}_{1,f_\varepsilon}^{-1}(u)\aleph(u)$ will vanish if $\boldsymbol{\xi}$ is observed exactly, and then the asymptotic result in Theorem 4 is the same as Theorem 3.1 obtained in Kai, Li and Zou (2011).

Using these initial estimators, we propose SCQR estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ and present its asymptotic normality. The SCQR estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are obtained by minimizing (3.5):

$$\begin{aligned}(\hat{\boldsymbol{\beta}}^\tau, \hat{\boldsymbol{\theta}}^\tau)^\tau& \\ &= \arg\min_{\boldsymbol{\beta}^*,\boldsymbol{\theta}^*} \sum_{i=1}^{n}\sum_{l=1}^{q} \rho_{\kappa_l}\big(Y_i - \hat{\boldsymbol{\xi}}_i^\tau \boldsymbol{\beta}^* - \mathbf{W}_i^\tau\boldsymbol{\theta}^* - \breve{\alpha}_{0l}(U_i) - \breve{\boldsymbol{\alpha}}(U_i)^\tau \mathbf{X}_i\big),\end{aligned} \tag{3.5}$$

with respect to $\boldsymbol{\beta}^*$ and $\boldsymbol{\theta}^*$.

Define $f_\varepsilon(c_\kappa) = (f_\varepsilon(c_{\kappa_1}), \dots, f_\varepsilon(c_{\kappa_q}))^\tau$, $c_{f_\varepsilon} = \sum_{l=1}^q f_\varepsilon(c_{\kappa_l})$. Moreover,

$$\boldsymbol{\lambda}_*(v) = E\big[(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau)^\tau | V = v\big],$$

$$\boldsymbol{\phi}_{f_\varepsilon}(u) = E\big[(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau)^\tau (\mathbf{0}_{(d+r)\times 1}^\tau, f_\varepsilon(c_\kappa)^\tau, c_{f_\varepsilon}\mathbf{X}^\tau)|U = u\big]\mathbf{A}_{1,f_\varepsilon}^{-1}(u).$$

Theorem 5. *Under the regularity conditions* (C1)–(C4), (C5)(ii), *and* (C7) *given in the Appendix, we have*

$$\sqrt{n}\begin{pmatrix}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \\ \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\end{pmatrix} \xrightarrow{\mathcal{L}} N\left(\mathbf{0}_{(d+r)\times 1}, \boldsymbol{\Lambda}_*^{-1}\boldsymbol{\Sigma}_{1*}\boldsymbol{\Lambda}_*^{-1} + \frac{1}{c_{f_\varepsilon}^2}\boldsymbol{\Lambda}_*^{-1}\boldsymbol{\Sigma}_{2*}\boldsymbol{\Lambda}_*^{-1}\right),$$

where $\boldsymbol{\Lambda}_* = E[\{(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau)^\tau\}^{\otimes 2}]$, $\boldsymbol{\Sigma}_{1*} = E[\boldsymbol{\lambda}_*^{\otimes 2}(V)]\boldsymbol{\beta}^\tau\boldsymbol{\Sigma}_{\mathbf{e}}\boldsymbol{\beta}$ *and*

$$\begin{aligned}\boldsymbol{\Sigma}_{2*} = \sum_{l=1}^{q}\sum_{s=1}^{q}(\kappa_l \wedge \kappa_s - \kappa_l\kappa_s)\big[&E\{\boldsymbol{\phi}_{f_\varepsilon}(U)\mathbf{M}_{[l]}(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau) + (\boldsymbol{\xi}^\tau, \mathbf{W}^\tau)^\tau \mathbf{M}_{[l]}^\tau \boldsymbol{\phi}_{f_\varepsilon}^\tau(U)\} \\ &+ E\{\boldsymbol{\phi}_{f_\varepsilon}(U)\mathbf{M}_{[l]}\mathbf{M}_{[s]}^\tau\boldsymbol{\phi}_{f_\varepsilon}^\tau(U)\} + \boldsymbol{\Lambda}_*\big].\end{aligned}$$

Remark 5. The first term $\mathbf{\Lambda}_*^{-1}\mathbf{\Sigma}_{1*}\mathbf{\Lambda}_*^{-1}$ in the asymptotic variance of Theorem 5 is due to the estimation of unobserved $\boldsymbol{\xi}$. If $\boldsymbol{\xi}$ is observed exactly (i.e., $\mathbf{e}=\mathbf{0}$), this term disappears analogous with Theorem 2. Moreover, if $\mathbf{e}=0$, $E(\mathbf{X}|U)=0$, $E(\boldsymbol{\xi}|U)=0$, and $E(\mathbf{W}|U)=0$, the asymptotic relative efficiency (ARE) of CQR for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ will be at least 86.4% when a large q is used, compared to the semi-least-squares estimate proposed by Li and Liang (2008). In the context of the measurement error considered in this paper, the ARE calculation for parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ involve an extra unknown term $\mathbf{\Sigma}_{1*}$, which involves the unknown covariance matrix $\mathbf{\Sigma}_{\mathbf{e}}$ of $\mathbf{e}$, unknown argument $\mathbf{\Lambda}_*(v)$ and unknown parameter $\boldsymbol{\beta}$. The ARE calculation for our models has no general result; however, it is of interest to explore solutions for different distributions of $\mathbf{e}$ and some particular structures of $\mathbf{\Lambda}_*(v)$.

Finally, we now refine the SCQR estimators of $\alpha_0(u)$ and $\boldsymbol{\alpha}(u)$ by using $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ obtained through (3.5). The refined estimators for $\alpha(u)$ and $\boldsymbol{\alpha}(u)$ are defined as

$$\begin{aligned}
&\big(\hat{\alpha}_{01}(u),\dots,\hat{\alpha}_{0q}(u),\hat{\alpha}_0'(u),\hat{\boldsymbol{\alpha}}^\tau(u),\hat{\boldsymbol{\alpha}}'^\tau(u)\big)^\tau \\
&\quad=\arg\min_{\substack{\alpha_{01},\dots,\alpha_{0q},\\ \alpha_0',\boldsymbol{\alpha},\boldsymbol{\alpha}'}}\sum_{i=1}^n\sum_{l=1}^q\rho_{\kappa_l}\big(Y_i-\hat{\boldsymbol{\xi}}_i^\tau\hat{\boldsymbol{\beta}}-\mathbf{W}_i^\tau\hat{\boldsymbol{\theta}} \\
&\qquad-\alpha_{0l}-\alpha_0'(U_i-u)-\boldsymbol{\alpha}^\tau\mathbf{X}_i-\boldsymbol{\alpha}'^\tau\mathbf{X}_i(U_i-u)\big) \\
&\qquad\times K_h(U_i-u).
\end{aligned}\tag{3.6}$$

The refined estimator for $\alpha_0(u)$ is further defined as

$$\hat{\alpha}_0(u)=\frac{1}{q}\sum_{l=1}^q\hat{\alpha}_{0l}(u).\tag{3.7}$$

We now study the asymptotic properties of $\hat{\alpha}_0(u)$ obtained from (3.7) and $\hat{\boldsymbol{\alpha}}(u)$ obtained from (3.6). Define

$$\begin{aligned}
\mathbf{A}_{2,f_\varepsilon}(u)&=E\left[\begin{matrix}\mathbf{C}_{f_\varepsilon} & f_\varepsilon(c_\kappa)\mathbf{X}^\tau\\ \mathbf{X}f_\varepsilon^\tau(c_\kappa) & c_{f_\varepsilon}\mathbf{X}^{\otimes 2}\end{matrix}\middle|U=u\right],\\
\mathbf{\Psi}_{f_\varepsilon}(u)&=E\big[\big(f_\varepsilon^\tau(c_\kappa),c_{f_\varepsilon}\mathbf{X}^\tau\big)^\tau\big(\boldsymbol{\xi}^{(2)}(V)\big)^\tau|U=u\big]\mathbf{C}\boldsymbol{\beta},\\
\mathbf{G}(u)&=\sum_{l=1}^q\sum_{s=1}^q(\kappa_l\wedge\kappa_s-\kappa_l\kappa_s)E\big[\big(m_l^\tau,\mathbf{X}^\tau\big)^\tau\big(m_s^\tau,\mathbf{X}^\tau\big)|U=u\big],
\end{aligned}$$

where $\mathbf{C}_{f_\varepsilon}=\mathrm{diag}(f_\varepsilon(c_{\kappa_1}),\dots,f_\varepsilon(c_{\kappa_q}))$. Moreover, let $\boldsymbol{\pi}_{q,1}=(\mathbf{1}_{q\times1}^\tau,\mathbf{0}_{p\times1}^\tau)^\tau$, $\boldsymbol{\pi}_{p,2}=(\mathbf{0}_{p\times q},\mathbf{I}_p)$, where $\mathbf{I}_p$ is an identical matrix of size p.

Theorem 6. *Under the regularity conditions* (C1)–(C4), (C5)(i), *and* (C7) *given in the Appendix, we have*

$$\begin{aligned}\sqrt{nh}\Bigg(&\hat{\alpha}_0(u) - \alpha_0(u) - \frac{1}{q}\sum_{l=1}^{q} c_{\kappa_l} - \frac{h^2\mu_{K_2}}{2}\alpha_0''(u) \\ &- \frac{h_o^2\mu_{L_2}}{2q}\boldsymbol{\pi}_{q,1}^{\tau}\mathbf{A}_{2,f_\varepsilon(u)}^{-1}\boldsymbol{\Psi}_{f_\varepsilon}(u)\Bigg) \\ &\xrightarrow{\mathcal{L}} N\bigg(0, \frac{\vartheta_{K_0}}{f_U(u)q^2}\boldsymbol{\pi}_{q,1}^{\tau}\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\mathbf{G}(u)\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\boldsymbol{\pi}_{q,1}\bigg).\end{aligned} \tag{3.8}$$

Moreover,

$$\begin{aligned}&\sqrt{nh}\bigg(\hat{\boldsymbol{\alpha}}(u) - \boldsymbol{\alpha}(u) - \frac{h^2\mu_{K_2}}{2}\boldsymbol{\alpha}''(u) - \frac{h_o^2\mu_{L_2}}{2}\boldsymbol{\pi}_{p,2}\mathbf{A}_{2,f_\varepsilon(u)}^{-1}\boldsymbol{\Psi}_{f_\varepsilon}(u)\bigg) \\ &\quad\xrightarrow{\mathcal{L}} N\bigg(\mathbf{0}_{p\times 1}, \frac{\vartheta_{K_0}}{f_U(u)}\boldsymbol{\pi}_{p,2}\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\mathbf{G}(u)\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\boldsymbol{\pi}_{p,2}^{\tau}\bigg).\end{aligned} \tag{3.9}$$

Remark 6. Note that our asymptotic variances in (3.9)–(3.9) are the same as Theorem 3.3 obtained in Kai, Li and Zou (2011). As indicated in Kai, Li and Zou (2011), CQR estimators utilize information shared across multiple quantile functions, which have competitive asymptotic efficiency compared with the least-squares methodology obtained by Zhou and Liang (2009), Li and Liang (2008). Moreover, the ARE is at least 88.9% for estimating varying coefficient functions (Kai, Li and Zou, 2011).

From (3.7), the baseline function estimator $\hat{\alpha}_0(u)$ converges to $\alpha_0(u)$ plus the average of the quantiles of error distribution; that is, $\frac{1}{q}\sum_{l=1}^{q} c_{\kappa_l}$. This average term is zero when the error distribution is symmetric. For non-asymmetric distributions, as q converges to infinity, this average converges to the mean of the error ε, which is also zero. For finite q, Sun, Gai and Lin (2013) designed weighted local linear composite quantile (WL-CQR) techniques to eliminate possible bias caused by the averages of the quantiles. However, deriving the asymptotic normality of the WL-CQR estimators for SQVCPLMeMs considered in this paper involves additional technicalities that go beyond the scope of the current paper. Therefore, this case will be considered in future research.

4 Numerical studies

In this section, we first conduct simulation studies to assess the performance of the proposed methods, and then we apply our methods to analyze a real dataset from a DMD study. We used the Epanechnikov kernel function $L(t) = K(t) = 0.75(1 - t^2)_+$ in the following numerical studies. For the estimators of

$\boldsymbol{\beta}_\kappa, \boldsymbol{\theta}_\kappa, \boldsymbol{\beta}$, and $\boldsymbol{\theta}$, as noted in Remark 2, the smoothing parameter h_k is chosen as $\hat{\sigma}_V n^{-1/3}$, where $\hat{\sigma}_V$ is the sample deviation of V. The smoothing parameter h is chosen as $\hat{\sigma}_U n^{-1/3}$, where $\hat{\sigma}_U$ is the sample deviation of U. These choices of h_k and h naturally meet the condition (C5)(ii). For the estimators of the nonparametric functions $\alpha_{0,\kappa}(u), \boldsymbol{\alpha}_\kappa(u), \alpha_0(u)$, and $\boldsymbol{\alpha}(u)$, Theorems 3 and 6 entail that the optimal theoretical bandwidth rate $h = Cn^{-1/5}$ is included in the condition (C5)(i). We suggest using the rule of thumb (Silverman, 1986) by choosing $h = \hat{\sigma}_U h^{-1/5}$ to meet condition (C5)(i).

4.1 A simulation study

Example. We generate 500 samples consisting of $n = 400$ observations from the following model

$$\begin{cases} Y = \boldsymbol{\beta}\boldsymbol{\xi} + \boldsymbol{\theta}^\tau \mathbf{W} + 2\cos^2(2\pi U) + \big(2\sin^2(2\pi U)\big)\mathbf{X} + \varepsilon, \\ \boldsymbol{\xi} = \boldsymbol{\xi}(V) \equiv 3V - \cos(V), \qquad \boldsymbol{\eta} = \boldsymbol{\xi}(V) + \mathbf{e}, \end{cases} \tag{4.1}$$

$\boldsymbol{\beta} = 2$, $\boldsymbol{\theta} = (3, 1.5, 2)$, $\mathbf{W} \sim N_3(\mathbf{0}, \Sigma_{\mathbf{W}})$, and $\Sigma_{\mathbf{W}} = (\sigma_{w,ij})$ with $\sigma_{w,ij} = 0.5^{|i-j|}$. $\mathbf{X} \sim N(0, 1)$, $U \sim \text{Uniform}[-1, 1]$, $V \sim N(0, 1)$ and is independent of $(U, \mathbf{W}, \mathbf{X})$, $\mathbf{e} \sim N(0, 1)$ and is independent of $(U, V, \mathbf{W}, \mathbf{X})$. The model error ε is independent with $(U, V, \mathbf{W}, \mathbf{X})$, and we consider three cases: (i) $\varepsilon \sim N(\mathbf{0}, 0.5^2)$. (ii) ε follows a t-distribution with 3 degrees of freedom. (iii) ε follows a mixture of normal distributions $0.9N(0, 1) + 0.1N(0, 10^2)$. Because of the independence condition between ε and $(U, V, \mathbf{W}, \mathbf{X})$, the SQR and SCQR procedures provide estimators for the same quantity and thus are directly comparable.

Performance of $\hat{\boldsymbol{\beta}}_\kappa, \hat{\boldsymbol{\theta}}_\kappa$ *and* $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}$. In Tables 1–3, we report the performances of the proposed estimators and the naive estimators (using $\boldsymbol{\eta}$ directly), and the simulation results for the benchmark estimators (i.e., all covariates are measured exactly) for $\boldsymbol{\beta}, \boldsymbol{\theta}$. The associated mean and associated standard errors of the estimators are also presented. We see that the estimated values of the SQR procedure and the SCQR procedure obtained by our proposed procedure, and the benchmark procedures are close to the true values in all three cases. This indicates our proposed method is promising. As anticipated, the naive estimator has severe bias and performs worse. This meets our expectation that large bias will occur if we ignore measurement errors. Moreover, the performance of the SQR procedure varies and depends heavily on the level of the quantile and the error distribution. Overall, SCQR outperforms SQR in general. From these numerical studies, the estimation procedure based on SCQR is very worthy of recommendation.

Performance of $\hat{\alpha}_{0,\kappa}(u), \hat{\boldsymbol{\alpha}}_\kappa(u)$ and $\hat{\alpha}_0(u), \hat{\boldsymbol{\alpha}}(u)$. Define the average square errors (ASEs) of a nonparametric estimator $\hat{\delta}(u)$ for its true value $\delta(u)$:

$$\text{ASE} = n_0^{-1} \sum_{i=1}^{n_0} \big\{\hat{\delta}(u_i) - \delta(u_i)\big\}^2,$$

Table 1 *The simulation results for normal distribution* $N(0, 0.5^2)$. *"MEAN" is the simulation mean; "SD" is the standard deviation. "P" stands for the proposed estimator, "B" stands for the benchmark estimator, and "N" stands for the naive estimator*

Method			$\theta_1 = 3$	$\theta_2 = 1.5$	$\theta_3 = 0.5$	$\beta_1 = 2$
$SCQR_9$	P	MEAN	2.9972	1.5012	0.4995	1.0522
		SD	0.0224	0.0242	0.0192	0.0380
	B	MEAN	2.9966	1.5014	0.5003	0.9998
		SD	0.0185	0.0206	0.0157	0.0041
	N	MEAN	2.9989	1.5133	0.5052	0.8295
		SD	0.0947	0.1171	0.0802	0.0275
$SQR_{0.05}$	P	MEAN	2.9982	1.5080	0.4963	1.0126
		SD	0.0393	0.0446	0.0389	0.0494
	B	MEAN	2.9978	1.5010	0.4977	0.9967
		SD	0.0310	0.0399	0.0312	0.0068
	N	MEAN	3.0061	1.5001	0.4897	0.6228
		SD	0.1372	0.1427	0.1128	0.0512
$SQR_{0.25}$	P	MEAN	3.0031	1.4983	0.5039	1.0343
		SD	0.0232	0.0239	0.0238	0.0440
	B	MEAN	3.0012	1.5013	0.5008	1.0011
		SD	0.0194	0.0214	0.0201	0.0066
	N	MEAN	2.9872	1.5092	0.4970	0.7322
		SD	0.0933	0.0853	0.0965	0.0319
$SQR_{0.5}$	P	MEAN	3.0006	1.5007	0.5021	1.0555
		SD	0.0219	0.0237	0.0215	0.0477
	B	MEAN	3.0017	1.5002	0.5011	1.0001
		SD	0.0172	0.0190	0.0197	0.0099
	N	MEAN	2.9953	1.4965	0.5057	0.8212
		SD	0.0907	0.0940	0.0923	0.0529
$SQR_{0.75}$	P	MEAN	2.9987	1.4986	0.5029	1.0700
		SD	0.0214	0.0260	0.0244	0.0453
	B	MEAN	2.9986	1.5025	0.5010	1.0003
		SD	0.0181	0.0192	0.0185	0.0062
	N	MEAN	3.0080	1.5081	0.4794	0.8962
		SD	0.1118	0.1339	0.1210	0.0294
$SQR_{0.95}$	P	MEAN	2.9973	1.5087	0.4998	1.0990
		SD	0.0385	0.0397	0.0370	0.0518
	B	MEAN	2.9979	1.5027	0.5033	1.0028
		SD	0.0315	0.0345	0.0309	0.0063
	N	MEAN	2.9584	1.5549	0.4933	0.9884
		SD	0.2575	0.2802	0.2359	0.0897

where $\{u_1, \ldots, u_{n_0}\}$ are the given grid points uniformly placed on $[0, 1]$ with $n_0 = 200$. We choose $\hat{\delta}(u_i)$ as the estimators for $\hat{\alpha}_{0,\kappa}(u_i), \hat{\boldsymbol{\alpha}}_\kappa(u_i)$, $\hat{\alpha}_0(u_i)$, and $\hat{\boldsymbol{\alpha}}(u_i)$, respectively. We evaluated the estimation procedures (2.10) and (3.6) for

Table 2 *The simulation results for $t(3)$ distribution. "MEAN" is the simulation mean; "SD" is the standard deviation. "P" stands for the proposed estimator, "B" stands for the benchmark estimator, and "N" stands for the naive estimator*

Method			$\theta_1 = 3$	$\theta_2 = 1.5$	$\theta_3 = 0.5$	$\beta_1 = 2$
$SCQR_9$	P	MEAN	3.0032	1.4984	0.5135	1.0458
		SD	0.0756	0.0708	0.0750	0.0476
	B	MEAN	3.0025	1.4983	0.5102	1.0006
		SD	0.0743	0.0689	0.0754	0.0155
	N	MEAN	3.0119	1.4917	0.5102	0.8362
		SD	0.1191	0.1219	0.1191	0.0305
$SQR_{0.05}$	P	MEAN	3.0015	1.4968	0.4819	1.0277
		SD	0.2236	0.2558	0.2405	0.0642
	B	MEAN	3.0009	1.4947	0.4892	0.9931
		SD	0.2233	0.2425	0.2285	0.0497
	N	MEAN	3.0290	1.5274	0.4467	0.7129
		SD	0.2707	0.3088	0.2649	0.0693
$SQR_{0.25}$	P	MEAN	3.0027	1.4926	0.5195	1.0454
		SD	0.0933	0.1210	0.1130	0.0559
	B	MEAN	2.9992	1.4946	0.5169	1.0026
		SD	0.0923	0.1166	0.1136	0.0233
	N	MEAN	2.9982	1.4950	0.5042	0.7742
		SD	0.1325	0.1630	0.1489	0.0429
$SQR_{0.5}$	P	MEAN	2.9919	1.4999	0.5175	1.0564
		SD	0.0911	0.1115	0.0852	0.0443
	B	MEAN	2.9955	1.5027	0.5145	1.0006
		SD	0.0907	0.1050	0.0827	0.0171
	N	MEAN	2.9735	1.5303	0.5060	0.8310
		SD	0.1221	0.1674	0.1369	0.0330
$SQR_{0.75}$	P	MEAN	3.0058	1.4970	0.5167	1.0693
		SD	0.0909	0.1182	0.1028	0.0403
	B	MEAN	3.0126	1.4923	0.5176	1.0003
		SD	0.0872	0.1186	0.1035	0.0194
	N	MEAN	3.0274	1.4834	0.5262	0.8972
		SD	0.1670	0.1926	0.1653	0.0399
$SQR_{0.95}$	P	MEAN	3.0184	1.4920	0.5148	1.0720
		SD	0.2381	0.3023	0.2697	0.0673
	B	MEAN	3.0082	1.5030	0.5055	0.9987
		SD	0.2441	0.3066	0.2728	0.0456
	N	MEAN	3.0392	1.5014	0.5031	0.9707
		SD	0.3101	0.4068	0.3214	0.0637

two scenarios: (i) using the estimated $\hat{\boldsymbol{\beta}}_\kappa, \hat{\boldsymbol{\theta}}_\kappa, \hat{\boldsymbol{\beta}},$ and $\hat{\boldsymbol{\theta}}$, (ii) using the true value $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$. We report the simulation means and standard derivations of the ASE for $\hat{\alpha}_{0,\kappa}(u_i)$, and $\hat{\boldsymbol{\alpha}}_\kappa(u_i)$, $\hat{\alpha}_0(u_i)$, and $\hat{\boldsymbol{\alpha}}(u_i)$ in Table 4 and Table 5. These results indicate that the performance of the benchmark estimators and the proposed estimators

Table 3 *The simulation results for mixture of normal distributions* $0.9N(0, 1) + 0.1N(0, 10^2)$. *"MEAN" is the simulation mean; "SD" is the standard deviation. "P" stands for the proposed estimator, "B" stands for the benchmark estimator, and "N" stands for the naive estimator*

Method			$\theta_1 = 3$	$\theta_2 = 1.5$	$\theta_3 = 0.5$	$\beta_1 = 2$
$SCQR_9$	P	MEAN	2.9830	1.5098	0.4906	1.0457
		SD	0.0747	0.0769	0.0719	0.0429
	B	MEAN	2.9874	1.5087	0.4921	0.9994
		SD	0.0730	0.0766	0.0730	0.0129
	N	MEAN	2.9832	1.5071	0.4998	0.8386
		SD	0.1227	0.1347	0.1144	0.0294
$SQR_{0.05}$	P	MEAN	2.9769	1.5063	0.5050	1.0212
		SD	0.1752	0.2040	0.1723	0.0569
	B	MEAN	2.9774	1.4916	0.5134	0.9962
		SD	0.1718	0.1997	0.1756	0.0373
	N	MEAN	2.9908	1.4820	0.5463	0.7073
		SD	0.2517	0.2648	0.2202	0.0608
$SQR_{0.25}$	P	MEAN	2.9943	1.5046	0.4910	1.0483
		SD	0.1197	0.1098	0.1144	0.0504
	B	MEAN	2.9986	1.5022	0.4923	0.9968
		SD	0.1143	0.1042	0.1095	0.0216
	N	MEAN	3.0082	1.4799	0.5064	0.7764
		SD	0.1608	0.1760	0.1561	0.0370
$SQR_{0.5}$	P	MEAN	2.9845	1.5076	0.4986	1.0509
		SD	0.1034	0.1288	0.0957	0.0490
	B	MEAN	2.9862	1.5032	0.4916	1.0006
		SD	0.1050	0.1278	0.0930	0.0191
	N	MEAN	2.9855	1.4888	0.4978	0.8353
		SD	0.1462	0.1527	0.1280	0.0306
$SQR_{0.75}$	P	MEAN	2.9965	1.4945	0.4868	1.0655
		SD	0.1254	0.1295	0.1000	0.0572
	B	MEAN	2.9914	1.4958	0.4891	1.0022
		SD	0.1258	0.1263	0.0995	0.0230
	N	MEAN	2.9908	1.5020	0.4699	0.8843
		SD	0.1662	0.1947	0.1405	0.0322
$SQR_{0.95}$	P	MEAN	3.0013	1.5003	0.5020	1.0767
		SD	0.1707	0.1801	0.1683	0.0650
	B	MEAN	3.0091	1.5074	0.4995	1.0008
		SD	0.1724	0.1790	0.1638	0.0348
	N	MEAN	2.9910	1.4891	0.5424	0.9733
		SD	0.3090	0.2961	0.3111	0.0596

works well regardless $\hat{\boldsymbol{\beta}}_\kappa, \hat{\boldsymbol{\theta}}_\kappa$, $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}$ or $\boldsymbol{\beta}, \boldsymbol{\theta}$ being used. This is not surprising because parameter estimators are always root-n consistent with higher convergence rates than nonparametric estimators. As a result, the benchmark estimator and the

Table 4 *The simulation results for* $2\cos^2(2\pi U)$ *with "MEAN"±"SD". "MEAN" is the simulation mean; "SD" is the standard deviation. "P" stands for the proposed estimator, "B" stands for the benchmark estimator, and "N" stands for the naive estimator*

		$2\cos^2(2\pi u)$		
Method		Normal distribution	$t(3)$ distribution	Mixture normal distribution
$SCQR_9$	P	0.0141 ± 0.0104	0.0645±0.0378	0.0477 ± 0.0282
	B	0.0081 ± 0.0013	0.0334 ± 0.0151	0.0272 ± 0.0118
	N	0.7627 ± 0.0987	0.7200 ± 0.1297	0.7156 ± 0.1287
$SQR_{0.05}$	P	0.0372 ± 0.0491	0.0913 ± 0.0413	0.0933 ± 0.0431
	B	0.0283 ± 0.0130	0.0889 ± 0.0309	0.0872 ± 0.0341
	N	1.2628 ± 0.3138	2.1133 ± 0.2356	1.2090 ± 0.1456
$SQR_{0.25}$	P	0.0535 ± 0.0242	0.0674 ± 0.0344	0.0649 ± 0.0403
	B	0.0448 ± 0.0074	0.0407 ± 0.0118	0.0460 ± 0.0211
	N	1.2021 ± 0.0996	0.3407 ± 0.0872	0.4519 ± 0.1227
$SQR_{0.50}$	P	0.0445 ± 0.0413	0.0569 ± 0.0341	0.0666 ± 0.0322
	B	0.0396 ± 0.0156	0.0346 ± 0.0124	0.0415 ± 0.0165
	N	1.1931 ± 0.1208	0.9576 ± 0.1789	0.8573 ± 0.1857
$SQR_{0.75}$	P	0.0312 ± 0.0218	0.0649 ± 0.0334	0.0702 ± 0.0379
	B	0.0150 ± 0.0123	0.0405 ± 0.0159	0.0504 ± 0.0159
	N	2.2695 ± 0.1481	1.6419 ± 0.2408	1.3760 ± 0.2091
$SQR_{0.95}$	P	0.0377 ± 0.0244	0.0987 ± 0.0456	0.0989 ± 0.0604
	B	0.0185 ± 0.0088	0.0820 ± 0.0378	0.0845 ± 0.0404
	N	3.5396 ± 0.1495	2.5556 ± 0.2322	1.9964 ± 0.3255

proposed estimator work satisfactorily under the two scenarios in term of the ASE. However, the naive procedure results in no-ignorable bias in the estimation of parameters $\boldsymbol{\theta}_\kappa$, $\boldsymbol{\theta}_\kappa$, $\boldsymbol{\beta}$, $\boldsymbol{\theta}$, and the nonparametric components $\alpha_0(u)$ and $\boldsymbol{\alpha}(u)$. The naive estimators by using true $\boldsymbol{\beta}$, $\boldsymbol{\theta}$ work well since no bias occurs. The estimators for $\alpha_0(u)$ and $\boldsymbol{\alpha}(u)$ with the estimated $\hat{\boldsymbol{\beta}}_\kappa, \hat{\boldsymbol{\theta}}_\kappa, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}$ perform as well as if we knew the true value of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ regardless the proposed estimation method or naive estimation. The performance of SQR varies and depends heavily on the level of quantiles and the error distributions, while SCQR outperforms more stable. Thus, the SCQR procedure outperforms the SQR procedure in this simulation.

4.2 An empirical example

We analyzed a dataset with 209 observations corresponding to blood samples from 192 patients from a DMD study. The patients were collected from a project to develop a screening program for female relatives of boys with DMD. The program's goal was to inform a woman of her chances of being a carrier based on serum markers, as well as her family pedigree. See Zhou and Liang (2009), Andrews and Herzberg (1985) for a detailed discussion of the dataset. In this dataset, the

Table 5 *The simulation results for* $2\sin^2(2\pi U)$ *with "MEAN"±"SD". "MEAN" is the simulation mean; "SD" is the standard deviation. "P" stands for the proposed estimator, "B" stands for the benchmark estimator, and "N" stands for the naive estimator*

Method		$2\sin^2(2\pi u)$		
		Normal distribution	$t(3)$ distribution	Mixture normal distribution
SCQR_9	P	0.0065 ± 0.0018	0.0261 ± 0.0133	0.0278 ± 0.0118
	B	0.0061 ± 0.0015	0.0257 ± 0.0122	0.0271 ± 0.0111
	N	0.0148 ± 0.0055	0.0360 ± 0.0216	0.0388 ± 0.0178
$\text{SQR}_{0.05}$	P	0.0141 ± 0.0104	0.0790 ± 0.0323	0.0972 ± 0.0410
	B	0.0083 ± 0.0027	0.0689 ± 0.0147	0.0938 ± 0.0393
	N	0.2280 ± 0.0710	0.1304 ± 0.0677	0.2322 ± 0.0899
$\text{SQR}_{0.25}$	P	0.0377 ± 0.0066	0.0378 ± 0.0089	0.0482 ± 0.0168
	B	0.0367 ± 0.0055	0.0375 ± 0.0088	0.0479 ± 0.0167
	N	0.0704 ± 0.0190	0.0650 ± 0.0270	0.0707 ± 0.0302
$\text{SQR}_{0.50}$	P	0.0256 ± 0.0144	0.0335 ± 0.0111	0.0396 ± 0.0149
	B	0.0246 ± 0.0094	0.0326 ± 0.0112	0.0381 ± 0.0145
	N	0.0433 ± 0.0113	0.0489 ± 0.0183	0.0532 ± 0.0223
$\text{SQR}_{0.75}$	P	0.0043 ± 0.0013	0.0379 ± 0.0162	0.0448 ± 0.0178
	B	0.0041 ± 0.0013	0.0366 ± 0.0156	0.0401 ± 0.0188
	N	0.0124 ± 0.0049	0.0438 ± 0.0198	0.0563 ± 0.0254
$\text{SQR}_{0.95}$	P	0.0114 ± 0.0066	0.0701 ± 0.0323	0.0960 ± 0.0320
	B	0.0084 ± 0.0028	0.0686 ± 0.0123	0.0929 ± 0.0327
	N	0.0132 ± 0.0049	0.0823 ± 0.0345	0.1168 ± 0.0444

serum marker creatine kinase (ck) is measured with errors and is inexpensive to obtain. We followed Zhou and Liang's (2009) procedure by regressing observed ck-$\boldsymbol{\eta}$ on the covariate U age. The covariate carrier status (cs-$\mathbf{W}$) and hemopexin-$\mathbf{X}$ are exactly measured. The marker lactate dehydrogenase (ld-Y) is very expensive to obtain, so it is of interest to predict the value ld-Y by using the level of ck, cs, hemopexin and age of the patients. Enzyme levels were measured in known carriers (75 observations) and in a group of non-carriers (134 observations). All the covariates are standardized. We used two SQR and SCQR procedures for this dataset.

The estimated values of the parameters for ck and cs are reported in Table 6. In Table 6, the 25% quantile, 50% quantile, and 75% quantile estimated values for ck are all negative, and those for cs are positive. Meanwhile, we use the 95%, 85%, 75%, 0.65%, 50%, 35%, 25%, 15%, and 5% quantiles for SCQR_9, the 85%, 75%, 0.65%, 50%, 35%, 25%, and 15% quantiles for SCQR_7, the 75%, 65%, 50%, 35%, and 25% quantiles for SCQR_5, and the 65%, 50%, and 35% quantiles for SCQR_3. The estimated SCQR values for ck are all negative, and those for cs are also positive. In Figures 1–2, we plot the 75%, 50%, and 25% quantile estimators for $\alpha_{0,\kappa}(u)$ and $\boldsymbol{\alpha}_\kappa(u)$ and the SCQR_7, SCQR_5, and

Table 6 *Analysis results for a DMD study*

	Method						
	$SQR_{0.25}$	$SQR_{0.50}$	$SQR_{0.75}$	$SCQR_9$	$SCQR_7$	$SCQR_5$	$SCQR_3$
Parameter estimate for ck	−1.2575	−1.1129	−0.3346	−1.0889	−1.0782	−1.5619	−1.2280
Parameter estimate for cs	0.3488	0.5872	0.7308	0.5612	0.5315	0.5213	0.5159

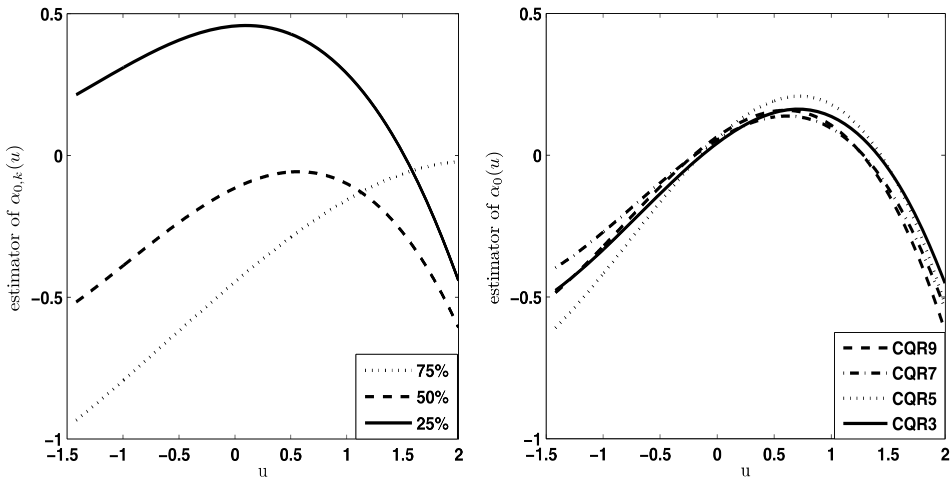

Figure 1 *The quantile estimators of $\hat{\alpha}_{0,\kappa}(\cdot)$ (the left panel) and SCQR estimators of $\alpha_0(\cdot)$ (the right panel).*

$SCQR_3$ estimators for $\alpha_0(u)$ and $\boldsymbol{\alpha}(u)$. The patterns of the $SCQR_7$, $SCQR_5$ and $SCQR_3$ estimators for $\alpha_0(u)$ and $\boldsymbol{\alpha}(u)$ are similar to the 50% quantile estimators for $\alpha_{0,\kappa}(u)$ and $\boldsymbol{\alpha}_\kappa(u)$, and are slightly different. The 75% quantile estimator for $\alpha_{0,\kappa}(u)$ increases as u increases, and the 25% quantile estimator for $\boldsymbol{\alpha}_\kappa(u)$ is different from the SCQR estimators.

Appendix

We present the conditions, prepare a preliminary lemma, and give the proofs of the main results.

A.1 Conditions

The following conditions are the regularity conditions for our asymptotic results.

(C1) The density function $f_V(v)$ of V is bounded away from 0 on $v \in \mathbb{V}$, where $\mathbb{V}$ is a bounded support of V. $f_V(v)$ and $\boldsymbol{\xi}(v)$ are twice continuously dif-

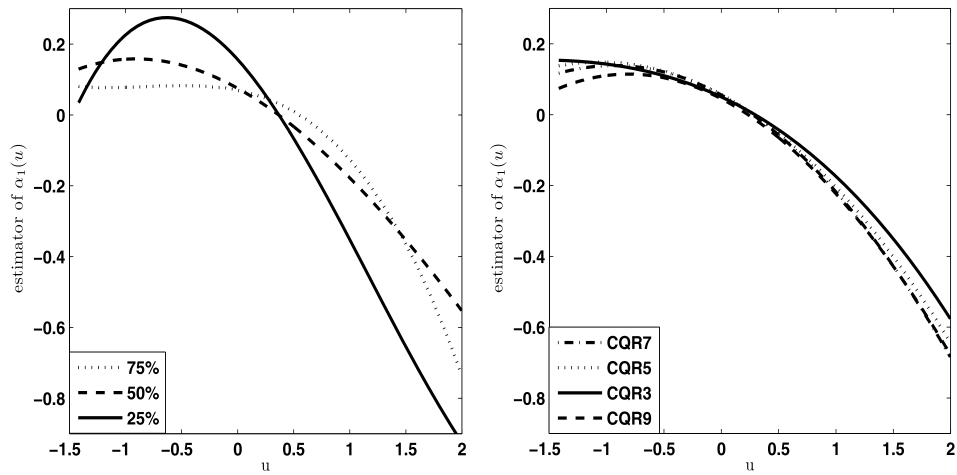

Figure 2 *The quantile estimators of* $\hat{\boldsymbol{\alpha}}_\kappa(\cdot)$ (*the left panel*) *and SCQR estimators of* $\boldsymbol{\alpha}(\cdot)$ (*the right panel*).

ferentiable with respect to v. Moreover, their second derivatives are uniformly Lipschitz continuous on $\mathbb{V}$: There exists a neighborhood of the origin, say $\mathcal{F}$, and a constant $c > 0$ such that for any $\varepsilon^* \in \mathcal{F}$, $|f_V^{(2)}(v)(v+\varepsilon^*) - f_V^{(2)}(v)(v)| < c|\varepsilon^*|$.

(C2) The random variable U has bounded support $\mathbb{U}$, and its density function $f_U(\cdot)$ is positive and has a continuous second derivative. Moreover, the joint density function $f_{U,V}(u,v)$ of (U,V) is continuous on the support $\mathbb{U} \times \mathbb{V}$.

(C3) The varying coefficients $\alpha_{0,\kappa}(u)$ and $\alpha_0(u)$ and the components of $\boldsymbol{\alpha}_\kappa(u)$ and $\boldsymbol{\alpha}(u)$ have a continuous second derivative in $u \in \mathbb{U}$.

(C4) The kernel functions $K(\cdot)$, $L(\cdot)$ are univariate bounded, continuous, and symmetric density functions satisfying that $\sup_t |K(t)| \le M_0$, $\sup_t |L(t)| \le M_0$ with a positive constant M_0, and $\int t^2 K(t)\,dt \ne 0$, $\int t^2 L(t)\,dt \ne 0$, and $\int |t|^j K(t)\,dt < \infty$, $\int |t|^j L(t)\,dt < \infty$. for $j = 1, 2, 3, 4$. Moreover, the second derivatives of $K(\cdot)$ and $L(\cdot)$ are bounded on $\mathbb{R}^1$.

(C5) The bandwidths h and h_k, $k = 1, \dots, d$ satisfy $h_k \asymp c_k h_o$ for some constant $c_k > 0$; $h \asymp c_h h_o$ for some constant $c_h > 0$. Moreover,
  (i) As $n \to \infty$, $h_o \to 0$ and $nh_o/(\log n) \to \infty$.
  (ii) As $n \to \infty$, $nh_o^4 \to 0$ and $nh_o^2/(\log n)^2 \to \infty$.

(C6) For the SQR procedure,
  (i) $F_\kappa(0|V = v, \mathbf{W} = \mathbf{w}, \mathbf{X} = \mathbf{x}, U = u) = \kappa$ for all $v, \mathbf{w}, \mathbf{x}, u$, and $f_\kappa(0| V = v, \mathbf{W} = \mathbf{w}, \mathbf{X} = \mathbf{x}, U = u) = \kappa$ is bounded away from zero and has a continuous and uniformly bounded derivative;
  (ii) $\mathbf{A}_1(u)$ defined in Theorem 1 and $\mathbf{A}_2(u)$ defined in Theorem 3 are nonsingular for all $u \in \mathbb{U}$. $\boldsymbol{\Lambda}_\kappa$ defined in Theorem 2 is a nonsingular matrix.

(C7) For the SCQR procedure,

(i) $f_\varepsilon(\cdot)$ is bounded away from zero and has a continuous and uniformly bounded derivative;

(ii) $\mathbf{A}_{1,f_\varepsilon}(u)$ defined in Theorem 4 and $\mathbf{A}_{2,f_\varepsilon}(u)$ defined in Theorem 6 are nonsingular for all $u \in \mathbb{U}$. $\mathbf{\Lambda}_*$ defined in Theorem 5 is a nonsingular matrix.

Conditions (C1)–(C3) are several mild smoothness conditions on the involved functions. $f_V(v)$ and $f_U(u)$ are all positive, which guarantees the denominators involved in the nonparametric smoothing are not equal to 0, once n is large enough. See, for example, Kai, Li and Zou (2010), Li and Liang (2008). Condition (C4) is the usual condition for the kernel functions $K(\cdot)$ and $L(\cdot)$. The Gaussian kernel and the quadratic kernel satisfy this condition. Condition (C5) is the condition for the bandwidths h and h_k in the nonparametric kernel smoothing. Conditions (C6)–(C7) are the regular conditions for the proposed SQR and SCQR procedures. For more details, see Kai, Li and Zou (2010).

A.2 A preliminary lemma

Lemma A.1. *Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d. random vectors, where the Y's are scalar random variables. Assume that $E|Y|^r < \infty$ and that $\sup_x \int |y|^r f(x, y)\,dy < \infty$, where f denotes the joint density of (X, Y). Let $K(\cdot)$ be a bounded positive function with bounded support, satisfying a Lipschitz condition. Then*

$$\begin{aligned}\sup_x\left|\frac{1}{n}\sum_{i=1}^n K_h(X_i - x)Y_i - E\big[K_h(X_i - x)Y_i\big]\right| \\ = O_P\big[\{nh/\log n\}^{-1/2}\big]\end{aligned} \tag{A.1}$$

provided that $n^{2\varepsilon-1}h \to \infty$ for some $\varepsilon < 1 - r^{-1}$.

Proof. Lemma A.1 follows a direct result of Mack and Silverman (1982). □

A.3 Proof of Theorem 1

In the following, for notational simplicity, let $r_i(u) = \alpha_{0\kappa}(U_i) - \alpha_{0\kappa}(u) - \alpha'_{0\kappa}(u)(U_i - u) + \mathbf{X}_i^\tau[\boldsymbol{\alpha}_\kappa(U_i) - \boldsymbol{\alpha}_\kappa(u) - \boldsymbol{\alpha}'_\kappa(u)(U_i - u)]$ and $K_h^u(U_i) = K_h(U_i - u)$. Define the local quantile estimators from (2.8) as

$$\begin{aligned}\breve{\boldsymbol{\delta}} = \sqrt{nh}\{&(\breve{\boldsymbol{\beta}}_\kappa - \boldsymbol{\beta}_\kappa)^\tau, (\breve{\boldsymbol{\theta}}_\kappa - \boldsymbol{\theta}_\kappa)^\tau, \\ &\breve{\alpha}_{0,\kappa}(u) - \alpha_{0,\kappa}(u), \big(\breve{\boldsymbol{\alpha}}_\kappa(u) - \boldsymbol{\alpha}_\kappa(u)\big)^\tau, \\ &h\big(\breve{\alpha}'_{0,\kappa}(u) - \alpha'_{0,\kappa}(u)\big), h\big(\breve{\boldsymbol{\alpha}}'_\kappa(u) - \boldsymbol{\alpha}'_\kappa(u)\big)^\tau\}.\end{aligned} \tag{A.2}$$

Proof. Recall that $\{\breve{\boldsymbol{\beta}}_\kappa, \breve{\boldsymbol{\theta}}_\kappa, \breve{\alpha}_{0\kappa}(u), \breve{\boldsymbol{\alpha}}_\kappa(u), \breve{\alpha}'_{0\kappa}(u), \breve{\boldsymbol{\alpha}}'_\kappa(u)\}$ minimizes

$$\sum_{i=1}^n \rho_\kappa\{Y_i - \boldsymbol{\beta}^\tau \hat{\boldsymbol{\xi}}_i - \boldsymbol{\theta}^\tau \mathbf{W}_i - \alpha_0 - \alpha_0'(U_i - u) - \boldsymbol{\alpha}^\tau \mathbf{X}_i - \boldsymbol{\alpha}'^\tau \mathbf{X}_i(U_i - u)\} K_h^u(U_i)$$

with respect to $\{\boldsymbol{\beta}, \boldsymbol{\theta}, \alpha_0, \alpha_0', \boldsymbol{\alpha}, \boldsymbol{\alpha}'\}$. We write $Y_i - \boldsymbol{\beta}^\tau \hat{\boldsymbol{\xi}}_i - \boldsymbol{\theta}^\tau \mathbf{W}_i - \alpha_0 - \alpha_0'(U_i - u) - \boldsymbol{\alpha}^\tau \mathbf{X}_i - \boldsymbol{\alpha}'^\tau \mathbf{X}_i(U_i - u) = \varepsilon_{\kappa,i} + r_i(u) - \hat{\Delta}_{\kappa,i} - \hat{\boldsymbol{\omega}}_{\kappa,i}$, where $\hat{\Delta}_{\kappa,i} = \tilde{\boldsymbol{\delta}}^\tau \hat{\mathbf{Z}}_i^u / \sqrt{nh}$, $\hat{\mathbf{Z}}_i^u = (\hat{\boldsymbol{\xi}}_i^\tau, \mathbf{W}_i^\tau, 1, \mathbf{X}_i^\tau, \frac{U_i - u}{h}, \frac{U_i - u}{h}\mathbf{X}_i^\tau)^\tau$ and $\hat{\boldsymbol{\omega}}_{\kappa,i} = \boldsymbol{\beta}_\kappa^\tau(\hat{\boldsymbol{\xi}}_i - \boldsymbol{\xi}_i)$. Then, $\hat{\boldsymbol{\delta}}$ defined in (A.2) is also the minimizer of

$$\mathcal{L}_n(\tilde{\boldsymbol{\delta}}) = \frac{1}{n}\sum_{i=1}^n \big[\rho_\kappa\{\varepsilon_{\kappa,i} + r_i(u) - \hat{\Delta}_{\kappa,i} - \hat{\boldsymbol{\omega}}_{\kappa,i}\} - \rho_\kappa\{\varepsilon_{\kappa,i} + r_i(u) - \hat{\boldsymbol{\omega}}_{\kappa,i}\}\big] K_h^u(U_i),$$

with respect to $\tilde{\boldsymbol{\delta}}$. By using the identity (Knight, 1998)

$$\rho_\kappa(x - y) - \rho_\kappa(x) = y\big[I\{x \le 0\} - \kappa\big] + \int_0^y \big[I\{x \le z\} - I\{x \le 0\}\big]\,dz, \quad \text{(A.3)}$$

we have

$$\begin{aligned}
\mathcal{L}_n(\tilde{\boldsymbol{\delta}}) &= \frac{1}{n}\sum_{i=1}^n \hat{\Delta}_{\kappa,i}\big[I\{\varepsilon_{\kappa,i} \le -r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i}\} - \kappa\big] K_h^u(U_i) \\
&\quad + \frac{1}{n}\sum_{i=1}^n \int_0^{\hat{\Delta}_{\kappa,i}} \big[I\{\varepsilon_{\kappa,i} \le -r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i} + z\} \\
&\quad - I\{\varepsilon_{\kappa,i} \le -r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i}\}\big]\,dz\, K_h^u(U_i) \\
&= \mathcal{L}_{n,1}(\tilde{\boldsymbol{\delta}}) + \mathcal{L}_{n,2}(\tilde{\boldsymbol{\delta}}).
\end{aligned}$$

Step 1. In this step, we analyze $\mathcal{L}_{n,2}(\tilde{\boldsymbol{\delta}})$. Define

$$\begin{aligned}
\mathcal{S}_n(u, t, \ell) &= \frac{1}{nh}\sum_{i=1}^n \int_0^\ell \big[I\{\varepsilon_{\kappa,i} \le -r_i(u) + t + z\} - I\{\varepsilon_{\kappa,i} \le -r_i(u) + t\} \\
&\quad - F_\kappa\big(-r_i(u) + t + z|\mathcal{O}_i\big) \\
&\quad + F_\kappa\big(-r_i(u) + t|\mathcal{O}_i\big)\big]\,dz\, K\big((U_i - u)/h\big) \\
&\overset{\text{def}}{=} \frac{1}{nh}\sum_{i=1}^n \mathcal{S}_n(u, t, \ell)[i].
\end{aligned} \quad \text{(A.4)}$$

Denote $\{\delta_1, \delta_2, \ldots, \delta_n\}$ be the Rademacher random variables; i.e., with $P(\delta_i = 1) = P(\delta_i = -1) = 1/2$ and independent with $\mathcal{R}_n = \{\mathcal{O}_i, \varepsilon_{\kappa,i}, i = 1, \ldots, n\}$, where $\mathcal{O}_i = \{V_i, \mathbf{W}_i, \mathbf{X}_i, U_i\}$. Let $\varsigma_n = \varsigma \frac{1}{nh}$, $\iota_n = c(h_o^2 + \sqrt{\frac{\log n}{nh_o}})$, here ς, c are two positive constants. The symmetrization Lemma (Pollard, 1984) implies that

$$\begin{aligned} & P\left(\sup_{|t|\le\iota_n,|\ell|\le\ell_0/\sqrt{nh}}\left|\sum_{i=1}^n \mathcal{S}_n(u,t,\ell)[i]\right|\ge nh\varsigma_n\right) \\ & \quad\le 4E\left\{P\left(\sup_{|t|\le\iota_n,|\ell|\le\ell_0/\sqrt{nh}}\left|\sum_{i=1}^n \delta_i\mathcal{S}_n(u,t,\ell)[i]\right|\ge nh\varsigma_n/4|\mathcal{R}_n\right)\right\}. \end{aligned} \tag{A.5}$$

Note that the class of functions $\mathcal{S}_n(u,t,\ell)[i]$ has envelope function $4M_0\ell_0/\sqrt{nh}$ by Condition (C4). Let $\mathcal{F}_n=\{\mathcal{S}_n(u,t,\ell)[i], i=1,\dots,n; |t|\le\iota_n, |\ell|\le\ell_0/\sqrt{nh}\}$ indexed by t,ℓ, and N^* be the smallest number such that $\min_{j_1,j_2\in\{1,\dots,N^*\}}\frac{1}{nh}\times\sum_{i=1}^n|\mathcal{S}_n(u,t,\ell)[i]-\mathcal{S}_n(u,t_{j_1},\ell_{j_2})[i]|\le\varsigma_n/8$, $\mathcal{S}_n(u,t_{j_1},\ell_{j_2})[i]\in\mathcal{F}_n$. Then, the conditional probability in (A.5) is further bounded by

$$\begin{aligned} & N^*P\left(\sup_{|t|\le\iota_n,|\ell|\le\ell_0/\sqrt{nh}}\left|\sum_{i=1}^n \delta_i\mathcal{S}_n(u,t_{j_1},\ell_{j_2})[i]\right|\ge nh\varsigma_n/4|\mathcal{R}_n\right) \\ & \quad\le N^*\max_{j_1\le N^*,j_2\le N^*}P\left(\left|\sum_{i=1}^n \delta_i\mathcal{S}_n(u,t_{j_1},\ell_{j_2})[i]\right|\ge nh\varsigma_n/8|\mathcal{R}_n\right) \\ & \quad\le 2N^*\max_{j_1\le N^*,j_2\le N^*}\exp\left\{-\frac{c_1^*n^2h^2\varsigma_n^2}{\sum_{i=1}^n\mathcal{S}_n^2(u,t_{j_1},\ell_{j_2})[i]}\right\}, \end{aligned} \tag{A.6}$$

for some positive constant c_1^*. The last inequality in (A.6) is asserted by the Hoeffding's Inequality. Note that $|\ell_{j_2}|\le\frac{\ell_0}{\sqrt{nh}}$, $|t_{j_1}|\le\iota_n$. Similar to the proof of Theorem 3.1 in Fan and Gijbels (1996), Taylor expansion for $F_\kappa(x|\mathcal{O}_i)$ around 0, we have

$$\frac{1}{nh}\sum_{i=1}^n E\left[E\{\mathcal{S}_n^2(u,t_{j_1},\ell_{j_2})[i]|\mathcal{O}_i\}\right]=O(\iota_n\ell_{j_2}^2). \tag{A.7}$$

Recalling that $\varsigma_n=\varsigma\frac{1}{nh}$, $\iota_n=c(h_o^2+\sqrt{\frac{\log n}{nh_o}})$ and $|\ell_{j_2}|\le\frac{\ell_0}{\sqrt{nh}}$, using (A.7), the last inequality in (A.6) is further bounded by $\exp\{-\frac{c_2^*n^2h^2\varsigma_n^2}{\iota_n\ell_{j_2}^2nh}\}=\exp\{-c_2^*\iota_n^{-1}\}$ for some constant c_2^* in probability. Moreover, note that the indicator function is a Vapnik–Chervonenkis class and conditional distribution function $F_\kappa(-r_i(u)+t|\mathcal{O}_i)$ is a Lipschitz function with respect to t, then the number N^* for the indicator functions and Lipschitz functions are bounded at most $\frac{k_0\max\{\iota_n,\frac{1}{\sqrt{nh}}\}}{\varsigma_n}$ for some positive constant k_0, see van der Vaart and Wellner (1996). As $h_o\to 0$, $\frac{nh_o}{\log n}\to\infty$, together with (A.5) and (A.6), we show that

$$\sup_{|t|\le\iota_n,|\ell|\le\ell_0/\sqrt{nh}}\left|\mathcal{S}_n(u,t,\ell)\right|=o_P\left(\frac{1}{nh}\right). \tag{A.8}$$

Directly using (A.8) and Taylor expansion, we have

$$\begin{aligned}\mathcal{L}_{n,2}(\tilde{\boldsymbol{\delta}}) &= \frac{1}{nh}\sum_{i=1}^{n}\int_0^{\hat{\Delta}_{\kappa,i}}\big[F_\kappa\big(-r_i(u)+\hat{\boldsymbol{\omega}}_{\kappa,i}+z|\mathcal{O}_i\big)-F_\kappa\big(-r_i(u)+\hat{\boldsymbol{\omega}}_{\kappa,i}|\mathcal{O}_i\big)\big]\,dz \\ &\quad\times K\big((U_i-u)/h\big)+o_P\Big(\frac{1}{nh}\Big) \\ &= \frac{1}{2nh}\tilde{\boldsymbol{\delta}}^\tau\Big[\frac{1}{n}\sum_{i=1}^{n}(\hat{\mathbf{Z}}_i^u)^{\otimes 2}f_\kappa\big(-r_i(u)+\hat{\boldsymbol{\omega}}_{\kappa,i}|\mathcal{O}_i\big)K_h^u(U_i)\Big]\tilde{\boldsymbol{\delta}}+o_P\Big(\frac{1}{nh}\Big) \\ &= \frac{1}{2nh}\big(f_U(u)\tilde{\boldsymbol{\delta}}^\tau\,\boldsymbol{\Xi}(u)\tilde{\boldsymbol{\delta}}+o_P(1)\big),\end{aligned} \tag{A.9}$$

where $\boldsymbol{\Xi}(u)=\mathrm{diag}(\mathbf{A}_1(u),\mu_{K_2}\mathbf{A}_2(u))$ is a quasi-diagonal matrix. Here $\mathbf{A}_1(u)$ and $\mathbf{A}_2(u)$ are defined as

$$\begin{aligned}\mathbf{A}_1(u) &= E\big[\mathbf{M}^{\otimes 2}f_\kappa(0|V,\mathbf{W},\mathbf{X},U)|U=u\big], \\ \mathbf{A}_2(u) &= E\big[f_\kappa(0|V,\mathbf{W},\mathbf{X},U)\big[(1,\mathbf{X}^\tau)^\tau\big]^{\otimes 2}|U=u\big],\end{aligned}$$

where $\mathbf{M}=(\boldsymbol{\xi}^\tau,\mathbf{W}^\tau,1,\mathbf{X}^\tau)^\tau$.

Step 2. In this step, we analyze $\mathcal{L}_{n,1}(\tilde{\boldsymbol{\delta}})$ and we have that

$$\begin{aligned}\mathcal{L}_{n,1}(\tilde{\boldsymbol{\delta}}) &= \frac{1}{n}\sum_{i=1}^{n}\big[I\{\varepsilon_{\kappa,i}\le -r_i(u)+\hat{\boldsymbol{\omega}}_{\kappa,i}\}-I\{\varepsilon_{\kappa,i}\le 0\}-F_\kappa\big(-r_i(u)+\hat{\boldsymbol{\omega}}_{\kappa,i}|\mathcal{O}_i\big) \\ &\quad+F_\kappa(0|\mathcal{O}_i)\big]\hat{\Delta}_{\kappa,i}K_h^u(U_i) \\ &\quad+\frac{1}{n}\sum_{i=1}^{n}\hat{\Delta}_{\kappa,i}\big[F_\kappa\big(-r_i(u)+\hat{\boldsymbol{\omega}}_{\kappa,i}|\mathcal{O}_i\big)-F_\kappa(0|\mathcal{O}_i)\big]K_h^u(U_i) \\ &\quad+\frac{1}{n}\sum_{i=1}^{n}\hat{\Delta}_{\kappa,i}\big[I\{\varepsilon_{\kappa,i}\le 0\}-\kappa\big]K_h^u(U_i) \\ &\stackrel{\text{def}}{=}\mathcal{L}_{n,1}^{[1]}(\tilde{\boldsymbol{\delta}})+\mathcal{L}_{n,1}^{[2]}(\tilde{\boldsymbol{\delta}})+\mathcal{L}_{n,1}^{[3]}(\tilde{\boldsymbol{\delta}}).\end{aligned}$$

Step 2.1. In this substep, we first deal with $\mathcal{L}_{n,1}^{[1]}(\tilde{\boldsymbol{\delta}})$. Let $\mathbf{Z}_i^u=(\boldsymbol{\xi}_i^\tau,\mathbf{W}_i^\tau,1,\mathbf{X}_i^\tau,\frac{U_i-u}{h},\frac{U_i-u}{h}\mathbf{X}_i^\tau)^\tau$ and $\mathbf{Z}_{i,s}^u$ be the s-component of $\mathbf{Z}_i^u$, $s=1,\ldots,(d+r+2p+2)$ and further we define

$$\begin{aligned}\upsilon_i(t,u) &= \big[I\{\varepsilon_{\kappa,i}\le -r_i(u)+t\}-I\{\varepsilon_{\kappa,i}\le 0\} \\ &\qquad -F_\kappa\big(-r_i(u)+t|\mathcal{O}_i\big)+F_\kappa(0|\mathcal{O}_i)\big]K\big((U_i-u)/h\big).\end{aligned}$$

We first show that $\sup_{|t|\le\iota_n}\frac{1}{\sqrt{nh}}|\sum_{i=1}^{n}\mathbf{Z}_{i,s}^u\upsilon_i(t,u)|=o_P(1)$ for $s=1,\ldots,(d+r+2p+2)$, where $\iota_n=c(h_o^2+\sqrt{\frac{\log n}{nh_o}})$ for some positive constant c. Similar to the analysis of (A.4)–(A.7), for any $\varsigma_0 > 0$, we have

$$
\begin{aligned}
&P\left(\sup_{|t|\le\iota_n}\frac{1}{\sqrt{nh}}\left|\sum_{i=1}^{n}\mathbf{Z}_{i,s}^{u}\upsilon_i(t,u)\right|\ge\varsigma_0\right)\\
&\quad\le 4E\left\{N^{**}\max_{1\le j\le N^{**}}P\left(\left|\sum_{i=1}^{n}\delta_i\mathbf{Z}_{i,s}^{u}\upsilon_i(t_j,u)\right|\ge\sqrt{nh}\varsigma_0/8\middle|\mathcal{R}_n\right)\right\} \qquad \text{(A.10)}\\
&\quad\le 4E\left\{N^{**}\max_{1\le j\le N^{**}}\exp\left(-\frac{\varsigma_0^*}{\frac{\sum_{i=1}^{n}(\mathbf{Z}_{i,s}^{u})^2\upsilon_i^2(t_j,u)}{nh}}\right)\right\},
\end{aligned}
$$

for some positive constant ς_0^*. As $|t_j|\le\iota_n$, Taylor expansion entails that $F_\kappa(t_j|\mathcal{O}_i)=F_\kappa(0|\mathcal{O}_i)+F_\kappa'(\tilde{t}_j|\mathcal{O}_i)t_j$, $|\tilde{t}_j|\le|t_j|$. Similar to the proof of Theorem 3.1 in Fan and Gijbels (1996), using Lemma A.1, it is seen that

$$
\begin{aligned}
&\frac{1}{nh}\sum_{i=1}^{n}E\big[E\{\upsilon_{i,s}^2(t_j,u)|\mathcal{O}_i\}\big]\\
&\quad=\frac{1}{nh}\sum_{i=1}^{n}E\big[(\mathbf{Z}_{i,s}^{u})^2K^2\big((U_i-u)/h\big)\\
&\qquad\times\{F_\kappa(t_j|\mathcal{O}_i)+F_\kappa(0|\mathcal{O}_i)-F_\kappa^2(t_j|\mathcal{O}_i)+F_\kappa^2(0|\mathcal{O}_i) \qquad \text{(A.11)}\\
&\qquad-2F_\kappa\big(\min\{t_j,0\}|\mathcal{O}_i\big)\}\big]\\
&\quad=O(\iota_n).
\end{aligned}
$$

Using (A.11), we have $\frac{\sum_{i=1}^{n}(\mathbf{Z}_{i,s}^{u})^2\upsilon_{i,s}^2(t_j,u)}{nh}=O_P(\iota_n)$. Similar to the analysis of (A.8), the number of N^{**} is bounded by $\frac{k_0^*\varsigma_0}{\iota_n}$ for some positive constant k_0^* (van der Vaart and Wellner, 1996). Then, the last inequality in (A.10) is bounded as $\frac{k_0^*\varsigma_0}{\iota_n}\exp\{-\varsigma_0^{**}\iota_n^{-1}\}$ in probability for some positive constant ς_0^{**}, and $\frac{k_0^*\varsigma_0}{\iota_n}\exp\{-\varsigma_0^{**}\iota_n^{-1}\}$ converges to 0 as $n\to\infty$. Thus,

$$
\frac{1}{nh}\tilde{\boldsymbol{\delta}}^{\tau}\left(\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\mathbf{Z}_i^{u}\upsilon_i(\hat{\boldsymbol{\omega}}_{\kappa,i},u)\right)=o_P\left(\frac{1}{nh}\right). \qquad \text{(A.12)}
$$

Moreover, let $\boldsymbol{\pi}=(\mathbf{I}_{d\times d},\mathbf{0}_{d\times(r+2p+2)})$ be a $d\times(d+r+2p+2)$ matrix, $\mathbf{C}=\mathrm{diag}(c_1^2,\dots,c_d^2)$, where c_i, $i=1,\dots,d$ are defined in Condition (C9). Similar to (A.12), using (2.5) and Lemma A.1,

$$
\begin{aligned}
&\frac{1}{nh}\tilde{\boldsymbol{\delta}}^{\tau}\left\{\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}(\hat{\mathbf{Z}}_i^{u}-\mathbf{Z}_i^{u})\upsilon_i(\hat{\boldsymbol{\omega}}_{\kappa,i},u)\right\}\\
&\quad=\frac{1}{nh}\tilde{\boldsymbol{\delta}}^{\tau}\left\{\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\left(\frac{h_o^2\mu_{L_2}\boldsymbol{\pi}^{\tau}}{2}\mathbf{C}\boldsymbol{\xi}^{(2)}(V_i)\right.\right. \qquad \text{(A.13)}
\end{aligned}
$$

$$+ O\left(\sqrt{\frac{\log n}{nh_o}}\right)\upsilon_i(\hat{\boldsymbol{\omega}}_{\kappa,i}, u)\bigg)\bigg\}$$

$$= o_P\left(\frac{1}{nh}\right).$$

Together with (A.12)–(A.13), we obtain that

$$\mathcal{L}_{n,1}^{[1]}(\tilde{\boldsymbol{\delta}}) = o_P\left(\frac{1}{nh}\right). \tag{A.14}$$

Step 2.2. In this substep, we analyze $\mathcal{L}_{n,1}^{[2]}(\tilde{\boldsymbol{\delta}})$ and $\mathcal{L}_{n,1}^{[3]}(\tilde{\boldsymbol{\delta}})$.

$$\begin{aligned}\mathcal{L}_{n,1}^{[2]}(\tilde{\boldsymbol{\delta}}) &= \frac{1}{nh}\tilde{\boldsymbol{\delta}}^\tau\left(\frac{1}{\sqrt{nh}}\sum_{i=1}^n \hat{\mathbf{Z}}_i^u f_\kappa(0|\mathcal{O}_i)\big(-r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i}\big)K\big((U_i - u)/h\big)\right) \\ &\quad + \frac{1}{nh}\tilde{\boldsymbol{\delta}}^\tau\left(\frac{1}{2\sqrt{nh}}\sum_{i=1}^n \hat{\mathbf{Z}}_i^u f_\kappa'(\tilde{l}_i|\mathcal{O}_i)\big(-r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i}\big)^2 K\big((U_i - u)/h\big)\right) \\ &\stackrel{\text{def}}{=} \mathcal{L}_{n,1}^{[2,1]}(\tilde{\boldsymbol{\delta}}) + \mathcal{L}_{n,1}^{[2,2]}(\tilde{\boldsymbol{\delta}}),\end{aligned}$$

where $\tilde{l}_i$ is between $-r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i}$ and zero. Using Lemma A.1, we have

$$\begin{aligned}&\frac{1}{nh}\sum_{i=1}^n \mathbf{Z}_i^u f_\kappa(0|\mathcal{O}_i) r_i(u) K\big((U_i - u)/h\big) \\ &\quad = h^2\left\{\frac{f_U(u)\mu_{K_2}}{2}\boldsymbol{\Xi}(u)\big(\mathbf{0}^\tau, \alpha_{0,\kappa}''(u), \boldsymbol{\alpha}_\kappa''^\tau(u), \mathbf{0}^\tau\big)^\tau + O_P\left(\sqrt{\frac{\log n}{nh}}\right)\right\}.\end{aligned} \tag{A.15}$$

Moreover, using the projection of U-statistics in Section 5.3.1 of Serfling (1980), we have

$$\begin{aligned}&\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \frac{\mathbf{Z}_i^u f_\kappa(0|\mathcal{O}_i)}{f_V(V_i)}\frac{1}{h}K\left(\frac{U_i - u}{h}\right)\frac{1}{h_s}L\left(\frac{V_j - V_i}{h_s}\right)\mathbf{e}_{j,s} \\ &\quad = \frac{1}{n}\sum_{i=1}^n \begin{pmatrix}\mathbf{m}(u, V_i) \\ \mathbf{0}\end{pmatrix}\frac{f_{U,V}(u, V_i)}{f_V(V_i)}\mathbf{e}_{i,s} + o_P\big(n^{-1/2}\big),\end{aligned} \tag{A.16}$$

where $\mathbf{m}(u, v) = E[\mathbf{M} f_\kappa(0|V, \mathbf{W}, \mathbf{X}, U)|U = u, V = v]$. Then, using (2.5) and Lemma A.1, together with (A.16) and Condition (C5), we have that

$$\begin{aligned}&\frac{1}{nh}\sum_{i=1}^n \mathbf{Z}_i^u f_\kappa(0|\mathcal{O}_i)\hat{\boldsymbol{\omega}}_{\kappa,i} K\big((U_i - u)/h\big) \\ &\quad = h_o^2\left\{\frac{\mu_{L_2} f_U(u)}{2}\big(\boldsymbol{\Upsilon}^\tau(u), \mathbf{0}^\tau\big)^\tau + O_P\left(\sqrt{\frac{\log n}{nh}}\right)\right\},\end{aligned} \tag{A.17}$$

where $\boldsymbol{\Upsilon}(u) = E[\mathbf{M}(\boldsymbol{\xi}^{(2)}(V))^\tau f_\kappa(0|V, \mathbf{W}, \mathbf{X}, U)|U = u]\mathbf{C}\boldsymbol{\beta}_\kappa$ and $\mathbf{C} = \operatorname{diag}(c_1^2, \ldots, c_d^2)$ and c_i's are defined in condition (C5). Similar to (A.15)–(A.17), we have $\mathcal{L}_{n,1}^{[2,2]}(\tilde{\boldsymbol{\delta}}) = o_P(\frac{1}{nh})$, and we obtain that

$$\begin{aligned}\mathcal{L}_{n,1}^{[3]}(\tilde{\boldsymbol{\delta}}) = {} & \frac{1}{nh}\tilde{\boldsymbol{\delta}}^\tau\left(\frac{1}{\sqrt{nh}}\sum_{i=1}^n \mathbf{Z}_i^u[I\{\varepsilon_{\kappa,i} \le 0\} - \kappa]K((U_i - u)/h)\right) \\ & + o_P\left(\frac{1}{nh}\right).\end{aligned} \tag{A.18}$$

Step 3. Together with asymptotic results (A.9) obtained in Step 1 and (A.14) and (A.18) in Step 2, we obtain that

$$\begin{aligned}\mathcal{L}_n(\tilde{\boldsymbol{\delta}}) = {} & \frac{1}{nh}\tilde{\boldsymbol{\delta}}^\tau\left(\frac{1}{2}f_U(u)\boldsymbol{\Xi}(u)\right)\tilde{\boldsymbol{\delta}} \\ & + \frac{1}{nh}\tilde{\boldsymbol{\delta}}^\tau\left(\frac{1}{\sqrt{nh}}\sum_{i=1}^n \mathbf{Z}_i^u[I\{\varepsilon_{\kappa,i} \le 0\} - \kappa]K((U_i - u)/h)\right) \\ & + \frac{1}{nh}\tilde{\boldsymbol{\delta}}^\tau\left(-\sqrt{nh}\frac{h^2\mu_{K_2}f_U(u)}{2}\boldsymbol{\Xi}(u)(\mathbf{0}^\tau, \alpha_{0,\kappa}''(u), \boldsymbol{\alpha}_\kappa''^\tau(u), \mathbf{0}^\tau)^\tau\right) \\ & + \frac{1}{nh}\tilde{\boldsymbol{\delta}}^\tau\left(\sqrt{nh}\frac{h_o^2\mu_{L_2}f_U(u)}{2}(\boldsymbol{\Upsilon}^\tau(u), \mathbf{0}^\tau)^\tau\right) + o_P\left(\frac{1}{nh}\right).\end{aligned} \tag{A.19}$$

By the convexity lemma (Pollard, 1991) and the quadratic approximation lemma (Fan and Gijbels, 1996), the minimizer of $\mathcal{L}_n(\tilde{\boldsymbol{\delta}})$ in (A.19) is expressed as

$$\begin{aligned}& \breve{\boldsymbol{\delta}} - \sqrt{nh}\frac{h^2\mu_{K_2}}{2}(\mathbf{0}^\tau, \alpha_{0,\kappa}''(u), \boldsymbol{\alpha}_\kappa''^\tau(u), \mathbf{0}^\tau)^\tau \\ & \quad + \sqrt{nh}\frac{h_o^2\mu_{L_2}}{2}\boldsymbol{\Xi}(u)^{-1}(\boldsymbol{\Upsilon}^\tau(u), \mathbf{0}^\tau)^\tau \\ & \quad = f_U^{-1}(u)\boldsymbol{\Xi}(u)^{-1}\left(\frac{1}{\sqrt{nh}}\sum_{i=1}^n \mathbf{Z}_i^u[I\{\varepsilon_{\kappa,i} \le 0\} - \kappa]K((U_i - u)/h)\right) \\ & \qquad + o_P(1).\end{aligned} \tag{A.20}$$

Recalling the definition of $\hat{\boldsymbol{\delta}}$, $\mathbf{Z}_i^u$, we complete the proof of Theorem 1. □

A.4 Proof of Theorem 2

Recall that $\hat{\boldsymbol{\beta}}_\kappa, \hat{\boldsymbol{\theta}}_\kappa$ minimize (A.21) with respect to $\boldsymbol{\beta}, \boldsymbol{\theta}$,

$$\sum_{i=1}^n \rho_\kappa\{Y_i - \boldsymbol{\beta}^\tau\hat{\boldsymbol{\xi}}_i - \boldsymbol{\theta}^\tau\mathbf{W}_i - \breve{\alpha}_{0,\kappa}(U_i) - \breve{\boldsymbol{\alpha}}^\tau(U_i)\mathbf{X}_i\}.$$

Define $\hat{\boldsymbol{\zeta}}_\kappa = \sqrt{n}((\hat{\boldsymbol{\beta}}_\kappa - \boldsymbol{\beta}_\kappa)^\tau, (\hat{\boldsymbol{\theta}}_\kappa - \boldsymbol{\theta}_\kappa)^\tau)$. Then $\hat{\boldsymbol{\zeta}}_\kappa$ is also the minimizer of (A.21) with respect to $\boldsymbol{\zeta}$:

$$\begin{aligned} \mathcal{Q}_n(\boldsymbol{\zeta}_\kappa) \stackrel{\text{def}}{=} \sum_{i=1}^n \big(\rho_\kappa\{\varepsilon_{\kappa,i} - n^{-1/2}\boldsymbol{\zeta}_\kappa^\tau \hat{\mathbf{S}}_i - \hat{\boldsymbol{\omega}}_{\kappa,i} - \tilde{r}(U_i)\} \\ - \rho_\kappa\{\varepsilon_{\kappa,i} - \hat{\boldsymbol{\omega}}_{\kappa,i} - \tilde{r}(U_i)\}\big), \end{aligned} \tag{A.21}$$

where $\hat{\mathbf{S}}_i = (\hat{\boldsymbol{\xi}}_i^\tau, \mathbf{W}_i^\tau)^\tau$ and $\tilde{r}(U_i) = (\tilde{\alpha}_{0,\kappa}(U_i) - \alpha_{0,\kappa}(U_i)) + (\tilde{\boldsymbol{\alpha}}_\kappa(U_i) - \boldsymbol{\alpha}_\kappa(U_i))^\tau \mathbf{X}_i$. Using (A.3) again, we have

$$\begin{aligned} \mathcal{Q}_n(\boldsymbol{\zeta}_\kappa) &= n^{-1/2}\boldsymbol{\zeta}_\kappa^\tau \sum_{i=1}^n \hat{\mathbf{S}}_i\big[I\{\varepsilon_{\kappa,i} \le \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)\} - \kappa\big] \\ &\quad + \sum_{i=1}^n \int_0^{n^{-1/2}\boldsymbol{\zeta}_\kappa^\tau \hat{\mathbf{S}}_i} \big[I\{\varepsilon_{\kappa,i} \le z + \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)\} - I\{\varepsilon_{\kappa,i} \le \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)\}\big]\,dz \\ &\stackrel{\text{def}}{=} \mathcal{Q}_{n,1}(\boldsymbol{\zeta}_\kappa) + \mathcal{Q}_{n,2}(\boldsymbol{\zeta}_\kappa). \end{aligned}$$

Step 1. Note that

$$\begin{aligned} \mathcal{Q}_{n,1}(\boldsymbol{\zeta}_\kappa) = n^{-1/2}\boldsymbol{\zeta}_\kappa^\tau \sum_{i=1}^n \hat{\mathbf{S}}_i\big[\{\varepsilon_{\kappa,i} \le \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)\} - I\{\varepsilon_{\kappa,i} \le 0\} \\ - F_\kappa\big(\hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big) + F_\kappa(0|\mathcal{O}_i)\big] \\ + n^{-1/2}\boldsymbol{\zeta}_\kappa^\tau \sum_{i=1}^n \hat{\mathbf{S}}_i\big[F_\kappa\big(\hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big) - F_\kappa(0|\mathcal{O}_i)\big] \\ + n^{-1/2}\boldsymbol{\zeta}_\kappa^\tau \sum_{i=1}^n \hat{\mathbf{S}}_i\big[I\{\varepsilon_{\kappa,i} \le 0\} - \kappa\big] \\ \stackrel{\text{def}}{=} \mathcal{Q}_{n,1}^{[1]}(\boldsymbol{\zeta}_\kappa) + \mathcal{Q}_{n,1}^{[2]}(\boldsymbol{\zeta}_\kappa) + \mathcal{Q}_{n,1}^{[3]}(\boldsymbol{\zeta}_\kappa). \end{aligned}$$

Let $\varphi_i(t) = I\{\varepsilon_{\kappa,i} \le t\} - I\{\varepsilon_{\kappa,i} \le 0\} - F_\kappa(t|\mathcal{O}_i) + F_\kappa(0|\mathcal{O}_i)$. Similar to the proof of (A.10), we have $\sup_{|t|\le \iota_n} n^{-1/2}|\sum_{i=1}^n \mathbf{S}_i\varphi_i(t)| = o_P(1)$ where $\mathbf{S}_i = (\boldsymbol{\xi}_i, \mathbf{W}_i)$. Using (2.5), (A.20) and Lemma A.1, it is easily seen that $\hat{\boldsymbol{\omega}}_i + \tilde{r}_i(U_i) = O_P(\iota_n)$. Then $\mathcal{Q}_{n,1}^{[1]}(\boldsymbol{\zeta}_\kappa) = o_P(1)\boldsymbol{\zeta}_\kappa$.

$$\begin{aligned} \mathcal{Q}_{n,1}^{[2]}(\boldsymbol{\zeta}_\kappa) &= n^{-1/2}\boldsymbol{\zeta}_\kappa^\tau \sum_{i=1}^n \hat{\mathbf{S}}_i f_\kappa(0|\mathcal{O}_i)[\hat{\boldsymbol{\omega}}_i + \tilde{r}_i(U_i)] \\ &\quad + n^{-1/2}\boldsymbol{\zeta}_\kappa^\tau \sum_{i=1}^n \hat{\mathbf{S}}_i f_\kappa'(\tilde{l}_i|\mathcal{O}_i)[\hat{\boldsymbol{\omega}}_i + \tilde{r}_i(U_i)]^2 \\ &= \mathcal{Q}_{n,1}^{[2,1]}(\boldsymbol{\zeta}_\kappa) + \mathcal{Q}_{n,1}^{[2,2]}(\boldsymbol{\zeta}_\kappa), \end{aligned}$$

where $\tilde{l}_i$ is between 0 and $\hat{\boldsymbol{\omega}}_i + \tilde{r}_i(U_i)$. Using (2.5), similar to (A.16), we have

$$
\begin{aligned}
n^{-1/2}&\sum_{i=1}^{n}\mathbf{S}_i f_\kappa(0|\mathcal{O}_i)\hat{\boldsymbol{\omega}}_i \\
&= \frac{n^{1/2}h_o^2\mu_{L_2}}{2}\frac{1}{n}\sum_{i=1}^{n}\mathbf{S}_i f_\kappa(0|\mathcal{O}_i)\boldsymbol{\xi}^{(2)}(V_i)^\tau \mathbf{C}\boldsymbol{\beta}_\kappa \\
&\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\big[\mathbf{S} f_\kappa(0|\mathcal{O})|V=V_i\big]\mathbf{e}_i^\tau\boldsymbol{\beta}_\kappa + o_P(1).
\end{aligned} \tag{A.22}
$$

Next, using (A.20), we have

$$
\begin{aligned}
\tilde{r}(U_i) = {}& \frac{\mu_{K_2}h^2}{2}\big(\alpha''_{0,\kappa}(U_i) + \boldsymbol{\alpha}''_\kappa(U_i)\mathbf{X}_i\big) - \frac{\mu_{L_2}h_o^2}{2}\big(\mathbf{0}^\tau, 1, \mathbf{X}_i^\tau\big)\mathbf{A}_1^{-1}(U_i)\boldsymbol{\Upsilon}(U_i) \\
& + f_U^{-1}(U_i)\big(\mathbf{0}^\tau, 1, \mathbf{X}_i^\tau\big) \\
& \times \mathbf{A}_1^{-1}(U_i)\left[\frac{1}{nh}\sum_{j=1}^{n}\mathbf{M}_j\big[I\{\varepsilon_{\kappa,j}\le 0\} - \kappa\big]K\left(\frac{U_j-U_i}{h}\right)\right] \\
& + o_P\left(\frac{1}{\sqrt{nh}}\right).
\end{aligned} \tag{A.23}
$$

where $\boldsymbol{\pi}_2 = (\mathbf{0}_{(p+1)\times(d+r)}, \mathbf{I}_{p+1})_{(d+r)\times(d+r+p+1)}$. Using the above expression (A.23), we have

$$
\begin{aligned}
n^{-1/2}&\sum_{i=1}^{n}\mathbf{S}_i f_\kappa(0|\mathcal{O}_i)\tilde{r}_i(U_i) \\
&= \frac{n^{1/2}h^2\mu_{K_2}}{2}\frac{1}{n}\sum_{i=1}^{n}\mathbf{S}_i f_\kappa(0|\mathcal{O}_i)\big(\alpha''_{0,\kappa}(U_i) + \boldsymbol{\alpha}''_\kappa(U_i)\mathbf{X}_i\big) \\
&\quad - \frac{n^{1/2}h_o^2\mu_{L_2}}{2}\frac{1}{n}\sum_{i=1}^{n}\mathbf{S}_i f_\kappa(0|\mathcal{O}_i)\big(\mathbf{0}^\tau, 1, \mathbf{X}_i^\tau\big)\mathbf{A}_1^{-1}(U_i)\boldsymbol{\Upsilon}(U_i) \\
&\quad + \frac{1}{n^{3/2}h}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\mathbf{S}_i f_\kappa(0|\mathcal{O}_i)}{f_U(U_i)}\big(\mathbf{0}^\tau, 1, \mathbf{X}_i^\tau\big) \\
&\quad \times \mathbf{A}_1^{-1}(U_i)\left[\mathbf{M}_j\big[I\{\varepsilon_{\kappa,j}\le 0\} - \kappa\big]K\left(\frac{U_j-U_i}{h}\right)\right] \\
&\quad + o_P(1).
\end{aligned} \tag{A.24}
$$

Using the projection of U-statistics in Section 5.3.1 of Serfling (1980), the third expression of (A.24) is asymptotically equivalent to

$$\begin{aligned}
&\frac{1}{n^{3/2}h}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\mathbf{S}_i f_\kappa(0|\mathcal{O}_i)}{f_U(U_i)}(\mathbf{0}^\tau,1,\mathbf{X}_i^\tau)\\
&\quad\times\mathbf{A}_1^{-1}(U_i)\Big[\mathbf{M}_j\big[I\{\varepsilon_{\kappa,j}\le 0\}-\kappa\big]K\Big(\frac{U_j-U_i}{h}\Big)\Big]\\
&\qquad=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\big[\mathbf{S}(\mathbf{0}^\tau,1,\mathbf{X}^\tau)f_\kappa(0|\mathcal{O})|U=U_i\big] \qquad\qquad (\text{A.25})\\
&\qquad\quad\times\mathbf{A}_1^{-1}(U_i)\mathbf{M}_i\big[I\{\varepsilon_{\kappa,i}\le 0\}-\kappa\big]\\
&\qquad\quad+o_P(1).
\end{aligned}$$

Using (A.23), (A.24) and (A.25), as $nh^4\to 0$, $nh_o^4\to 0$, we have

$$\begin{aligned}
\mathcal{Q}_{n,1}^{[2,1]}(\boldsymbol{\zeta}_\kappa)&=\boldsymbol{\zeta}_\kappa^\tau\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\big[\mathbf{S}f_\kappa(0|\mathcal{O})|V=V_i\big]\mathbf{e}_i^\tau\boldsymbol{\beta}_\kappa\right\}\\
&\quad+\boldsymbol{\zeta}_\kappa^\tau\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}E\big[\mathbf{S}(\mathbf{0}^\tau,1,\mathbf{X}^\tau)f_\kappa(0|\mathcal{O})|U=U_i\big]\right. \qquad (\text{A.26})\\
&\quad\left.\times\mathbf{A}_1^{-1}(U_i)\mathbf{M}_i\big[I\{\varepsilon_{\kappa,i}\le 0\}-\kappa\big]+o_P(1)\right\}.
\end{aligned}$$

Using (2.5) and (A.23), as $\frac{(\log n)^2}{nh^2}\to 0$, it is easily seen that

$$\begin{aligned}
&n^{-1/2}\sum_{i=1}^{n}\hat{\mathbf{S}}_i f_\kappa'(\tilde{l}_i|\mathcal{O}_i)\big[\hat{\boldsymbol{\omega}}_i+\tilde{r}_i(U_i)\big]^2\\
&\qquad=n^{1/2}O_P\Big(h_o^4+h^4+\frac{\log n}{nh}\Big) \qquad\qquad (\text{A.27})\\
&\qquad=o_P(1).
\end{aligned}$$

Together with (A.26) and (A.27), the asymptotic expression of $\mathcal{Q}_{n,1}^{[2]}(\boldsymbol{\zeta}_\kappa)$ is obtained. Using (2.5) again, we have

$$\mathcal{Q}_{n,1}^{[3]}(\boldsymbol{\zeta}_\kappa)=\boldsymbol{\zeta}_\kappa^\tau\left\{n^{-1/2}\sum_{i=1}^{n}\mathbf{S}_i\big[I\{\varepsilon_{\kappa,i}\le 0\}-\kappa\big]+o_P(1)\right\}. \qquad (\text{A.28})$$

As a result, the asymptotic expression of $\mathcal{Q}_{n,1}(\boldsymbol{\zeta}_\kappa)$ is the summation of (A.26) and (A.28).

Step 2. In this step, we analyze $\mathcal{Q}_{n,2}(\boldsymbol{\zeta}_\kappa)$.

$$\begin{aligned}\mathcal{Q}_{n,2}(\boldsymbol{\zeta}_\kappa) &= \sum_{i=1}^n \int_0^{n^{-1/2}\zeta_\kappa^\tau \hat{\mathbf{S}}_i} \big[I\{\varepsilon_{\kappa,i} \le z + \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)\} - I\{\varepsilon_{\kappa,i} \le \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)\} \\ &\quad - F_\kappa\big(z + \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big) - F_\kappa\big(\hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big)\big]\,dz \\ &\quad + \sum_{i=1}^n \int_0^{n^{-1/2}\zeta_\kappa^\tau \hat{\mathbf{S}}_i} \big[F_\kappa\big(z + \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big) - F_\kappa\big(\hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big)\big]\,dz \\ &\stackrel{\text{def}}{=} \mathcal{Q}_{n,2}^{[1]}(\boldsymbol{\zeta}_\kappa) + \mathcal{Q}_{n,2}^{[2]}(\boldsymbol{\zeta}_\kappa).\end{aligned}$$

Similar to (A.8) and (A.9), we have $\mathcal{Q}_{n,2}^{[1]}(\boldsymbol{\zeta}_\kappa) = o_P(1)$. Next, Taylor expansion entails that

$$\begin{aligned}\mathcal{Q}_{n,2}^{[2]}(\boldsymbol{\zeta}_\kappa) &= \sum_{i=1}^n \int_0^{n^{-1/2}\zeta_\kappa^\tau \hat{\mathbf{S}}_i} \big[F_\kappa\big(z + \hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big) - F_\kappa\big(\hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big)\big]\,dz \\ &= \frac{1}{2}\boldsymbol{\zeta}_\kappa^\tau \left[\frac{1}{n}\sum_{i=1}^n (\hat{\mathbf{S}}_i)^{\otimes 2} f_\kappa\big(\hat{\boldsymbol{\omega}}_i + \tilde{r}(U_i)|\mathcal{O}_i\big)\right]\boldsymbol{\zeta}_\kappa + o_P(1) \qquad \text{(A.29)} \\ &= \frac{1}{2}\boldsymbol{\zeta}_\kappa^\tau \{E[\mathbf{S}^{\otimes 2} f_\kappa(0|\mathcal{O})]\}\boldsymbol{\zeta}_\kappa + o_P(1) \stackrel{\text{def}}{=} \frac{1}{2}\boldsymbol{\zeta}_\kappa^\tau \boldsymbol{\Lambda}_\kappa \boldsymbol{\zeta}_\kappa + o_P(1).\end{aligned}$$

Together with (A.26), (A.28) and (A.29), using the convexity lemma (Pollard, 1991) and the quadratic approximation lemma (Fan and Gijbels, 1996), the minimizer of $\mathcal{Q}_n(\boldsymbol{\zeta})$ in (A.21) is expressed as

$$\begin{aligned}\hat{\boldsymbol{\zeta}}_\kappa &= \boldsymbol{\Lambda}_\kappa^{-1}\Bigg\{\frac{1}{\sqrt{n}}\sum_{i=1}^n E[\mathbf{S} f_\kappa(0|\mathcal{O})|V = V_i]\mathbf{e}_i^\tau \boldsymbol{\beta}_\kappa \\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^n [\mathbf{S}_i + E[\mathbf{S}(\mathbf{0}^\tau, 1, \mathbf{X}^\tau) f_\kappa(0|\mathcal{O})|U = U_i]\mathbf{A}_1^{-1}(U_i)\mathbf{M}_i] \qquad \text{(A.30)} \\ &\quad \times \big[I\{\varepsilon_{\kappa,i} \le 0\} - \kappa\big]\Bigg\} \\ &\quad + o_P(1).\end{aligned}$$

Recalling the definition of $\hat{\boldsymbol{\zeta}}_\kappa$, $\mathbf{S}_i$, we complete the proof of Theorem 2.

A.5 Proof of Theorem 3

The proof of Theorem 3 is similar to the proof of Theorem 1. We present some main steps. Let $\mathbf{N}_i^u = (1, \mathbf{X}_i^\tau, \frac{U_i - u}{h}, \mathbf{X}_i^\tau \frac{U_i - u}{h})^\tau$ and

$$\begin{aligned}\hat{\varkappa} = \sqrt{nh}\big(&(\hat{\alpha}_{0,\kappa}(u) - \alpha_{0,\kappa}(u)), (\hat{\boldsymbol{\alpha}}_\kappa^\tau(u) - \boldsymbol{\alpha}_{0,\kappa}^\tau(u))^\tau, \\ &h(\hat{\alpha}'_{0,\kappa}(u) - \alpha'_{0,\kappa}(u)), h(\hat{\boldsymbol{\alpha}}'^\tau_{0,\kappa}(u) - \boldsymbol{\alpha}'^\tau_{0,\kappa}(u))^\tau\big).\end{aligned} \qquad \text{(A.31)}$$

Then (A.31) is the minimizer of (A.32) with respect to $\varkappa$,

$$\mathcal{F}_n(\varkappa) \stackrel{\text{def}}{=} \frac{1}{nh}\sum_{i=1}^{n}\big(\rho_\kappa\{\varepsilon_{\kappa,i} - n^{-1/2}\hat{\boldsymbol{\zeta}}_\kappa^\tau \hat{\mathbf{S}}_i - \hat{\boldsymbol{\omega}}_{\kappa,i} + r_i(u) - \varkappa^\tau \mathbf{N}_i^u/\sqrt{nh}\} \\ - \rho_\kappa\{\varepsilon_{\kappa,i} - n^{-1/2}\hat{\boldsymbol{\zeta}}_\kappa^\tau \hat{\mathbf{S}}_i - \hat{\boldsymbol{\omega}}_{\kappa,i} + r_i(u)\}\big)K\big((U_i - u)/h\big). \tag{A.32}$$

Note that (A.32) is decomposed as

$$\begin{aligned}\mathcal{F}_n(\varkappa) &= \frac{1}{nh}\varkappa^\tau\Bigg(\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\mathbf{N}_i^u\big[I\{\varepsilon_{\kappa,i} \le -r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i} + n^{-1/2}\hat{\boldsymbol{\zeta}}_\kappa^\tau\hat{\mathbf{S}}_i\} - \kappa\big] \\ &\quad \times K\big((U_i - u)/h\big)\Bigg) \\ &\quad + \frac{1}{nh}\Bigg(\sum_{i=1}^{n}\int_0^{\varkappa^\tau \mathbf{N}_i^u/\sqrt{nh}}\big[I\{\varepsilon_{\kappa,i} \le -r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i} + n^{-1/2}\hat{\boldsymbol{\zeta}}_\kappa^\tau\hat{\mathbf{S}}_i + z\} \\ &\quad - I\{\varepsilon_{\kappa,i} \le -r_i(u) + \hat{\boldsymbol{\omega}}_{\kappa,i} + n^{-1/2}\hat{\boldsymbol{\zeta}}_\kappa^\tau\hat{\mathbf{S}}_i\}\big]\,dz\Bigg)K\big((U_i - u)/h\big) \\ &= \mathcal{F}_{n,1}(\varkappa) + \mathcal{F}_{n,2}(\varkappa).\end{aligned}$$

Similar to the analysis in the Step 2.1 of Theorem 1, we have

$$\begin{aligned}\mathcal{F}_{n,1}(\varkappa) &= \frac{1}{nh}\varkappa^\tau\left(-\frac{\sqrt{nh}h^2\mu_{K_2}f_U(u)}{2}\operatorname{diag}\big(\mathbf{A}_2(u), \mathbf{0}_{l_1\times l_1}\big)\begin{pmatrix}\alpha_{0,\kappa}''(u)\\ \boldsymbol{\alpha}_\kappa''(u)\\ \mathbf{0}\end{pmatrix}_{l_2\times 1}\right) \\ &\quad + \frac{1}{nh}\varkappa^\tau\left(\frac{\sqrt{nh}h_o^2\mu_{L_2}f_U(u)}{2}\begin{pmatrix}\boldsymbol{\Psi}(u)\\ \mathbf{0}\end{pmatrix}_{l_2\times 1}\right) \\ &\quad + \frac{1}{nh}\varkappa^\tau\left(\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\mathbf{N}_i^u[I\{\varepsilon_{\kappa,i} \le 0\} - \kappa]K\big((U_i - u)/h\big)\right) \\ &\quad + o_P\left(\frac{1}{nh}\right),\end{aligned} \tag{A.33}$$

where $l_1 = p+1, l_2 = 2p+2$, $\boldsymbol{\Psi}(u) = E[(1, \mathbf{X}^\tau)^\tau(\boldsymbol{\xi}^{(2)}(V))^\tau f_\kappa(0|V, \mathbf{W}, \mathbf{X}, U)| U = u]\mathbf{C}\boldsymbol{\beta}_\kappa$. Similar to the analysis in the Step 1 of Theorem 1, we have

$$\mathcal{F}_{n,2}(\varkappa) = \frac{1}{nh}\varkappa^\tau\{f_U(u)\operatorname{diag}\big(\mathbf{A}_2(u), \mu_{K_2}\mathbf{A}_2(u)\big)\}\varkappa \\ + o_P\left(\frac{1}{nh}\right). \tag{A.34}$$

Together with (A.34) and (A.34), using the convexity lemma (Pollard, 1991) and the quadratic approximation lemma (Fan and Gijbels, 1996), we have

$$\begin{aligned}
&\sqrt{nh}\Big((\hat{\alpha}_{0,\kappa}-\alpha_{0,\kappa}), (\hat{\boldsymbol{\alpha}}_\kappa^\tau-\boldsymbol{\alpha}_\kappa^\tau)^\tau - \frac{\mu_{K_2}h^2}{2}\big(\alpha''_{0,\kappa}(u), \boldsymbol{\alpha}''^\tau_\kappa(u)\big)^\tau \\
&\quad + \frac{\mu_{L_2}h_o^2}{2}\mathbf{A}_2^{-1}(u)\boldsymbol{\Psi}(u)\Big) \\
&\quad = \mathbf{A}_2^{-1}(u) f_U^{-1}(u)\left\{\frac{1}{\sqrt{nh}}\sum_{i=1}^n (1,\mathbf{X}_i^\tau)^\tau [I\{\varepsilon_{\kappa,i}\le 0\}-\kappa] K\big((U_i-u)/h\big)\right\} \\
&\qquad + o_P(1).
\end{aligned} \tag{A.35}$$

We complete the proof of Theorem 3.

A.6 Proof of Theorem 4

In the following, $\vec{r}_i(u)=\alpha_0(U_i)-\alpha_0(u)-\alpha_0'(u)(U_i-u)+\mathbf{X}_i^\tau[\boldsymbol{\alpha}(U_i)-\boldsymbol{\alpha}(u)-\boldsymbol{\alpha}'(u)(U_i-u)]$. Furthermore, let

$$\begin{aligned}
\breve{\boldsymbol{\nu}} = \sqrt{nh}\{&(\breve{\boldsymbol{\beta}}-\boldsymbol{\beta})^\tau, (\breve{\boldsymbol{\theta}}-\boldsymbol{\theta})^\tau, \\
&\breve{\alpha}_{01}(u)-\alpha_0(u)-c_{\kappa_1},\ldots,\breve{\alpha}_{0q}(u)-\alpha_0(u)-c_{\kappa_q}, \\
&\big(\breve{\boldsymbol{\alpha}}(u)-\boldsymbol{\alpha}(u)\big)^\tau, h\big(\breve{\alpha}'_0(u)-\alpha_0'(u)\big), h\big(\breve{\boldsymbol{\alpha}}'(u)-\boldsymbol{\alpha}'(u)\big)^\tau\}.
\end{aligned} \tag{A.36}$$

and we define $\hat{\mathbf{Z}}_{i,l}^u=(\hat{\boldsymbol{\xi}}_i^\tau, \mathbf{W}_i^\tau, m_l^\tau, \mathbf{X}_i^\tau, \frac{U_i-u}{h}, \frac{U_i-u}{h}\mathbf{X}_i^\tau)^\tau$ where m_l is a q-vector with 1 at the lth position and 0 elsewhere. Recall that $\{\breve{\boldsymbol{\beta}}, \breve{\boldsymbol{\theta}}, \breve{\alpha}_{0l}(u), \breve{\boldsymbol{\alpha}}(u), \breve{\alpha}'_0(u), \breve{\boldsymbol{\alpha}}'(u)\}$ minimizes

$$\begin{aligned}
&\sum_{l=1}^n\sum_{i=1}^n \rho_{\kappa_l}\{Y_i-\boldsymbol{\beta}^\tau\hat{\boldsymbol{\xi}}_i-\boldsymbol{\theta}^\tau\mathbf{W}_i-\alpha_{0l}-\alpha_0'(U_i-u)-\boldsymbol{\alpha}^\tau\mathbf{X}_i-\boldsymbol{\alpha}'^\tau\mathbf{X}_i(U_i-u)\} \\
&\quad\times K_h^u(U_i)
\end{aligned}$$

with respect to $\{\boldsymbol{\beta},\boldsymbol{\theta},\alpha_{0l},\alpha_0',\boldsymbol{\alpha},\boldsymbol{\alpha}'\}$. We write $Y_i-\boldsymbol{\beta}^\tau\hat{\boldsymbol{\xi}}_i-\boldsymbol{\theta}^\tau\mathbf{W}_i-\alpha_{0l}-\alpha_0'(U_i-u)-\boldsymbol{\alpha}^\tau\mathbf{X}_i-\boldsymbol{\alpha}'^\tau\mathbf{X}_i(U_i-u)=\varepsilon_{\kappa,i}-c_{\kappa_l}+\vec{r}_i(u)-\hat{\Delta}_{i,l}-\hat{\boldsymbol{\omega}}_i$, where $\hat{\Delta}_{i,l}=\breve{\boldsymbol{\nu}}^\tau\hat{\mathbf{Z}}_{i,l}^u/\sqrt{nh}$, $\hat{\boldsymbol{\omega}}_i=\boldsymbol{\beta}^\tau(\hat{\boldsymbol{\xi}}_i-\boldsymbol{\xi}_i)$. Then, $\breve{\boldsymbol{\nu}}$ is also the minimizer of

$$\begin{aligned}
\mathcal{L}_n^*(\boldsymbol{\nu}) = \frac{1}{n}\sum_{l=1}^q\sum_{i=1}^n &\{\rho_{\kappa_l}\big(\varepsilon_i-c_{\kappa_l}+\vec{r}_i(u)-\hat{\Delta}_{i,l}-\hat{\boldsymbol{\omega}}_i\big) \\
&-\rho_{\kappa_l}\big(\varepsilon_i-c_{\kappa_l}+\vec{r}_i(u)-\hat{\boldsymbol{\omega}}_i\big)\} \\
&\times K_h^u(U_i)
\end{aligned}$$

with respect to $\boldsymbol{\nu}$. Using (A.3), we have

$$\begin{aligned}\mathcal{L}_n^*(\boldsymbol{\nu}) &= \frac{1}{n}\sum_{l=1}^{q}\sum_{i=1}^{n}\hat{\Delta}_{i,l}\big[I\{\varepsilon_i \le c_{\kappa_l} - \vec{r}_i(u) + \hat{\boldsymbol{\omega}}_i\} - \kappa_l\big]K_h^u(U_i)\\ &\quad + \frac{1}{n}\sum_{l=1}^{q}\sum_{i=1}^{n}\int_0^{\hat{\Delta}_{i,l}}\big[I\{\varepsilon_i \le c_{\kappa_l} - \vec{r}_i(u) + \hat{\boldsymbol{\omega}}_i + z\}\\ &\quad - I\{\varepsilon_i \le c_{\kappa_l} - \vec{r}_i(u) + \hat{\boldsymbol{\omega}}_i\}\big]\,dz K_h^u(U_i)\\ &= \mathcal{L}_{n,1}^*(\boldsymbol{\nu}) + \mathcal{L}_{n,2}^*(\boldsymbol{\nu}).\end{aligned}$$

Similar to the proof of $\mathcal{L}_{n,2}(\boldsymbol{\delta})$ in the Step 1 of Theorem 1, it is shown that

$$\begin{aligned}\mathcal{L}_{n,2}^*(\boldsymbol{\nu}) &= \frac{1}{nh}\boldsymbol{\nu}^\tau\sum_{l=1}^{q}\left[\frac{1}{2n}\sum_{i=1}^{n} f_\varepsilon\big(c_{\kappa_l} - \vec{r}_i(u) + \hat{\boldsymbol{\omega}}_i\big)K_h^u(U_i)\big(\hat{\mathbf{Z}}_{i,l}^u\big)^{\otimes 2}\right]\boldsymbol{\nu}\\ &\quad + o_P\left(\frac{1}{nh}\right) \qquad\qquad (\text{A.37})\\ &= \frac{1}{nh}\boldsymbol{\nu}^\tau\left[\frac{1}{2}f_U(u)\boldsymbol{\Xi}_{f_\varepsilon}(u)\right]\boldsymbol{\nu} + o_P\left(\frac{1}{nh}\right),\end{aligned}$$

where $\boldsymbol{\Xi}_{f_\varepsilon}(u) = \operatorname{diag}(\mathbf{A}_{1,f_\varepsilon}(u), c_{f_\varepsilon}\mu_{K_2}\mathbf{A}_2(u))$ is a quasi-diagonal matrix. Here $c_{f_\varepsilon} = \sum_{l=1}^{q} f_\varepsilon(c_{\kappa_l})$, and $\mathbf{A}_{1,f_\varepsilon}(u)$ is defined as

$$\mathbf{A}_{1,f_\varepsilon}(u) = \sum_{l=1}^{q} f_\varepsilon(c_{\kappa_l})E\big[\mathbf{M}_l^{\otimes 2}|U=u\big], \qquad \mathbf{M}_{[l]} = \big(\boldsymbol{\xi}^\tau, \mathbf{W}^\tau, m_l^\tau, \mathbf{X}^\tau\big)^\tau.$$

Similar to the proof of $\mathcal{L}_{n,1}(\boldsymbol{\delta})$ in the Step 2 of Theorem 1, it is shown that

$$\begin{aligned}\mathcal{L}_{n,1}^*(\boldsymbol{\nu}) &= \frac{\boldsymbol{\nu}^\tau}{nh}\left(\frac{1}{\sqrt{nh}}\sum_{l=1}^{q}\sum_{i=1}^{n}\mathbf{Z}_{i,l}^u\big[I\{\varepsilon_i \le c_{\kappa_l}\} - \kappa_l\big]K\big((U_i - u)/h\big)\right)\\ &\quad + \frac{\boldsymbol{\nu}^\tau}{nh}\Bigg(-\frac{\sqrt{nh}\mu_{K_2}f_U(u)h^2}{2}\\ &\quad \times\sum_{l=1}^{q} f_\varepsilon(c_{\kappa_l})\mathbf{T}_l(u)\big(\mathbf{0}^\tau, m_l^\tau\alpha_0''(u), \boldsymbol{\alpha}''^\tau(u), \mathbf{0}^\tau\big)^\tau\Bigg) \qquad\qquad (\text{A.38})\\ &\quad + \frac{\boldsymbol{\nu}^\tau}{nh}\left(-\frac{\sqrt{nh}\mu_{L_2}f_U(u)h_o^2}{2}\big(\aleph^\tau(u), \mathbf{0}^\tau\big)^\tau\right) + o_P\left(\frac{1}{nh}\right),\end{aligned}$$

where $\mathbf{Z}_{i,l}^u = (\mathbf{M}_{[l]i}^\tau, \frac{U_i-u}{h}, \frac{U_i-u}{h}\mathbf{X}_i^\tau)^\tau$, $\mathbf{T}_l(u) = E[\{(\mathbf{M}_{[l]}^\tau, 1, \mathbf{X}^\tau)^\tau\}^{\otimes 2}|U=u]$ with $\mathbf{M}_{[l]i} = (\boldsymbol{\xi}_i^\tau, \mathbf{W}_i^\tau, m_l^\tau, \mathbf{X}_i^\tau)^\tau$ and $\mathbf{M}_{[l]i} = (\boldsymbol{\xi}^\tau, \mathbf{W}^\tau, m_l^\tau, \mathbf{X}^\tau)^\tau$, and

$$\aleph(u) = \sum_{l=1}^{q} f_\varepsilon(c_{\kappa_l})E\big[\mathbf{M}_{[l]}\big(\boldsymbol{\xi}^{(2)}(V)\big)^\tau|U=u\big]\mathbf{C}\boldsymbol{\beta}.$$

Together with (A.38) and (A.39), the proof of Theorem 4 is completed by directly using the convexity lemma (Pollard, 1991), the quadratic approximation lemma (Fan and Gijbels, 1996) and proofs of Theorem 3.1 of Kai, Li and Zou (2011).

A.7 Proof of Theorem 5

The proof of Theorem 5 is similar to the proof of Theorem 2, we only present the main steps. Define $\hat{\boldsymbol{\zeta}} = \sqrt{n}((\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^\tau, (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^\tau)^\tau$. Recalling the definition of $\hat{\mathbf{S}}_i$ used in the proof of Theorem 2, then the estimator $\hat{\boldsymbol{\zeta}}$ is also the minimizer of

$$\begin{aligned}\mathcal{H}_n(\boldsymbol{\zeta}) \stackrel{\text{def}}{=} \sum_{i=1}^n \sum_{l=1}^q \big(\rho_{\kappa_l}\{\varepsilon_i - c_{\kappa_l} - n^{-1/2}\boldsymbol{\zeta}^\tau \hat{\mathbf{S}}_i - \hat{\boldsymbol{\omega}}_i - \vec{r}_l(U_i)\} \\ - \rho_\kappa\{\varepsilon_i - c_{\kappa_l} - \hat{\boldsymbol{\omega}}_i - \vec{r}_l(U_i)\}\big),\end{aligned}$$

where $\vec{r}_l(U_i) = (\breve{\alpha}_{0l}(U_i) - \alpha_0(U_i) - c_{\kappa_l}) + (\breve{\boldsymbol{\alpha}}(U_i) - \boldsymbol{\alpha}(U_i))^\tau \mathbf{X}_i$. We have

$$\begin{aligned}\mathcal{H}_n(\boldsymbol{\zeta}) &= n^{-1/2}\boldsymbol{\zeta}^\tau \sum_{i=1}^n \sum_{l=1}^q \hat{\mathbf{S}}_i\big[I\{\varepsilon_i \le c_{\kappa_l} + \hat{\boldsymbol{\omega}}_i + \vec{r}_l(U_i)\} - \kappa_l\big] \\ &\quad + \sum_{i=1}^n \sum_{l=1}^q \int_0^{n^{-1/2}\boldsymbol{\zeta}^\tau \hat{\mathbf{S}}_i} \big[I\{\varepsilon_i \le z + c_{\kappa_l} + \hat{\boldsymbol{\omega}}_i + \vec{r}_l(U_i)\} \\ &\quad - I\{\varepsilon_i \le c_{\kappa_l} + \hat{\boldsymbol{\omega}}_i + \vec{r}_l(U_i)\}\big]\,dz \\ &\stackrel{\text{def}}{=} \mathcal{H}_{n,1}(\boldsymbol{\zeta}) + \mathcal{H}_{n,2}(\boldsymbol{\zeta}).\end{aligned}$$

Let $\mathbf{X}_{[l]i} = (\mathbf{0}^\tau_{(d+r)\times 1}, m_l^\tau, \mathbf{X}_i^\tau)^\tau$ in the following. Similar to (A.25), the asymptotic expression (A.39) and the projection of U-statistics in Section 5.3.1 of Serfling (1980) entail that

$$\begin{aligned}&\frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{l=1}^q f_\varepsilon(c_{\kappa_l}) \mathbf{S}_i \mathbf{X}^\tau_{[l]i} \frac{\mathbf{A}^{-1}_{1,f_\varepsilon}(U_i)}{f_U(U_i)} \\ &\quad \times \left[\frac{1}{nh} \sum_{s=1}^n \sum_{j=1}^q \mathbf{M}_{[j]s}\big[I\{\varepsilon_s \le c_{\kappa_j}\} - \kappa_j\big] K\left(\frac{U_s - U_i}{h}\right)\right] \\ &= \sum_{l=1}^q \sum_{j=1}^q f_\varepsilon(c_{\kappa_l}) \left[\frac{1}{n^{3/2}h} \sum_{i=1}^n \sum_{s=1}^n \mathbf{S}_i \mathbf{X}^\tau_{[l]i} \frac{\mathbf{A}^{-1}_{1,f_\varepsilon}(U_i)}{f_U(U_i)} \right. \\ &\quad \left. \times \mathbf{M}_{[j]s}\big[I\{\varepsilon_s \le c_{\kappa_j}\} - \kappa_j\big] K\left(\frac{U_s - U_i}{h}\right)\right] \\ &= \sum_{l=1}^q \sum_{j=1}^q f_\varepsilon(c_{\kappa_l}) \frac{1}{\sqrt{n}} \sum_{i=1}^n E\big[\mathbf{S}\mathbf{X}^\tau_{[l]} | U = U_i\big] \\ &\quad \times \mathbf{A}^{-1}_{1,f_\varepsilon}(U_i) \mathbf{M}_{[j]i}\big[I\{\varepsilon_i \le c_{\kappa_j}\} - \kappa_j\big].\end{aligned} \tag{A.39}$$

Similar to the Step 1 in the proof of Theorem 2, we have

$$\begin{aligned}
&\mathcal{H}_{n,1}(\boldsymbol{\zeta}) \\
&= \boldsymbol{\zeta}^\tau \Bigg[\sum_{l=1}^{q}\sum_{j=1}^{q} f_\varepsilon(c_{\kappa_l})\frac{1}{\sqrt{n}}\sum_{i=1}^{n} E\big[\mathbf{S}\mathbf{X}_{[l]}^\tau | U = U_i\big] \\
&\quad \times \mathbf{A}_{1,f_\varepsilon}^{-1}(U_i)\mathbf{M}_{[j]i}\big[I\{\varepsilon_i \le c_{\kappa_j}\} - \kappa_j\big]\Bigg] \qquad (A.40)\\
&\quad + \boldsymbol{\zeta}^\tau \Bigg[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{l=1}^{q}\{\mathbf{S}_i\big[I\{\varepsilon_i \le c_{\kappa_l}\} - \kappa_l\big] \\
&\quad + f_\varepsilon(c_{\kappa_l})E[\mathbf{S}|V = V_i]\mathbf{e}_i^\tau \boldsymbol{\beta}\} + o_P(1)\Bigg]
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{H}_{n,2}(\boldsymbol{\zeta}) &= \boldsymbol{\zeta}^\tau \left[\frac{1}{2}\sum_{l=1}^{q} f_\varepsilon(c_{\kappa_l})\frac{1}{n}\sum_{i=1}^{n}\hat{\mathbf{S}}_i^{\otimes 2}\right]\boldsymbol{\zeta} \\
&= \boldsymbol{\zeta}^\tau \left[\frac{1}{2}\sum_{l=1}^{q} f_\varepsilon(c_{\kappa_l})E\big(\mathbf{S}^{\otimes 2}\big) + o_P(1)\right]\boldsymbol{\zeta}.
\end{aligned} \qquad (A.41)$$

Using (A.40) and (A.41), the proof of Theorem 4 is completed by directly using the convexity lemma (Pollard, 1991), the quadratic approximation lemma (Fan and Gijbels, 1996). We complete the proof of Theorem 5.

A.8 Proof of Theorem 6

The proof of Theorem 6 is similar to the proof of Theorem 3. We present some main steps. Let $\mathbf{N}_{li}^u = (m_l^\tau, \mathbf{X}_i^\tau, \frac{U_i-u}{h}, \mathbf{X}_i^\tau \frac{U_i-u}{h})^\tau$ and

$$\begin{aligned}
\hat{\boldsymbol{\varphi}} = \sqrt{nh}\big(&(\hat{\alpha}_{0l}(u) - \alpha_0(u) - c_{\kappa_1}), \ldots, \\
&(\hat{\alpha}_{0q}(u) - \alpha_0(u) - c_{\kappa_q}), \big(\hat{\boldsymbol{\alpha}}^\tau(u) - \boldsymbol{\alpha}^\tau(u)\big)^\tau, \\
&h\big(\hat{\alpha}'_0(u) - \alpha'_0(u)\big), h\big(\hat{\boldsymbol{\alpha}}'^\tau(u) - \boldsymbol{\alpha}'^\tau(u)\big)^\tau\big).
\end{aligned} \qquad (A.42)$$

Recalling the definition of $\hat{\boldsymbol{\zeta}}$ used in the proof of Theorem 5, Then (A.42) is the minimizer of (A.43) with respect to $\boldsymbol{\varphi}$,

$$\begin{aligned}
\mathcal{D}_n(\boldsymbol{\varphi}) \overset{\text{def}}{=} \frac{1}{nh}\sum_{l=1}^{q}\sum_{i=1}^{n}\big(&\rho_{\kappa_l}\{\varepsilon_i - c_{\kappa_l} - n^{-1/2}\hat{\boldsymbol{\zeta}}^\tau\hat{\mathbf{S}}_i - \hat{\boldsymbol{\omega}}_i + \vec{r}_i(u) - \boldsymbol{\varphi}^\tau\mathbf{N}_{li}^u/\sqrt{nh}\} \\
&- \rho_{\kappa_l}\{\varepsilon_i - c_{\kappa_l} - n^{-1/2}\hat{\boldsymbol{\zeta}}^\tau\hat{\mathbf{S}}_i - \hat{\boldsymbol{\omega}}_i + \vec{r}_i(u)\}\big)K\big((U_i - u)/h\big).
\end{aligned} \qquad (A.43)$$

Note that (A.43) is decomposed as

$$
\begin{aligned}
&\mathcal{D}_n(\boldsymbol{\varphi})\\
&\quad= \frac{1}{nh}\boldsymbol{\varphi}^\tau\Bigg(\frac{1}{\sqrt{nh}}\sum_{l=1}^{q}\sum_{i=1}^{n}\mathbf{N}_{li}^u\big[I\{\varepsilon_i \le c_{\kappa_l} - \vec{r}_i(u) + \hat{\boldsymbol{\omega}}_i + n^{-1/2}\hat{\boldsymbol{\zeta}}^\tau\hat{\mathbf{S}}_i\} - \kappa_l\big]\\
&\qquad \times K\Big(\frac{U_i-u}{h}\Big)\Bigg)\\
&\qquad + \frac{1}{nh}\Bigg(\sum_{l=1}^{q}\sum_{i=1}^{n}\int_0^{\boldsymbol{\varphi}^\tau\mathbf{N}_{li}^u/\sqrt{nh}}\big[I\{\varepsilon_i \le c_{\kappa_l} - \vec{r}_i(u) + \hat{\boldsymbol{\omega}}_i + n^{-1/2}\hat{\boldsymbol{\zeta}}^\tau\hat{\mathbf{S}}_i + z\}\\
&\qquad - I\{\varepsilon_i \le c_{\kappa_l} - \vec{r}_i(u) + \hat{\boldsymbol{\omega}}_i + n^{-1/2}\hat{\boldsymbol{\zeta}}^\tau\hat{\mathbf{S}}_i\}\big]\,dz\Bigg)K\big((U_i-u)/h\big)\\
&\quad= \mathcal{D}_{n,1}(\boldsymbol{\varphi}) + \mathcal{D}_{n,2}(\boldsymbol{\varphi}).
\end{aligned}
$$

Similar to the analysis in the Step 1 of Theorem 1, we have

$$
\begin{aligned}
\mathcal{D}_{n,2}(\boldsymbol{\varphi}) &= \frac{f_U(u)}{nh}\\
&\quad \times \boldsymbol{\varphi}^\tau\left[\operatorname{diag}\left\{E\left(\begin{matrix}\mathbf{C}_{f_\varepsilon} & f_\varepsilon(c_\kappa)\mathbf{X}^\tau\\ \mathbf{X}f_\varepsilon^\tau(c_\kappa) & c_{f_\varepsilon}\mathbf{X}^{\otimes 2}\end{matrix}\,\middle|\, U=u\right), c_{f_\varepsilon}\mu_{K_2}\mathbf{A}_2(u)\right\}\right]\boldsymbol{\varphi}\\
&\quad + o_P\Big(\frac{1}{nh}\Big)\\
&\stackrel{\text{def}}{=} \frac{1}{nh}\boldsymbol{\varphi}^\tau\Big[\frac{1}{2}f_U(u)\operatorname{diag}\big(\mathbf{A}_{2,f_\varepsilon}(u), c_{f_\varepsilon}\mu_{K_2}\mathbf{A}_2(u)\big)\Big]\boldsymbol{\varphi} + o_P\Big(\frac{1}{nh}\Big),
\end{aligned}
\tag{A.44}
$$

where $\mathbf{C}_{f_\varepsilon} = \operatorname{diag}(f_\varepsilon(c_{\kappa_1}), \dots, f_\varepsilon(c_{\kappa_q}))$. Moreover, similar to the analysis in the Step 2 of Theorem 1, we have

$$
\begin{aligned}
\mathcal{D}_{n,1}(\boldsymbol{\varphi}) &= \frac{1}{nh}\boldsymbol{\varphi}^\tau\left(-\frac{\sqrt{nh}h^2\mu_{K_2}f_U(u)}{2}\operatorname{diag}\big(\mathbf{A}_{2,f_\varepsilon}(u), \mathbf{0}_{l_1\times l_1}\big)\begin{pmatrix}\alpha_0''(u)1_q\\ \boldsymbol{\alpha}''(u)\\ \mathbf{0}\end{pmatrix}_{l_2^*\times 1}\right)\\
&\quad + \frac{1}{nh}\boldsymbol{\varphi}^\tau\left(\frac{\sqrt{nh}h_o^2\mu_{L_2}f_U(u)}{2}\begin{pmatrix}\boldsymbol{\Psi}_{f_\varepsilon}(u)\\ \mathbf{0}\end{pmatrix}_{l_2\times 1}\right)\\
&\quad + \frac{1}{nh}\boldsymbol{\varphi}^\tau\left(\frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\sum_{l=1}^{q}\mathbf{N}_{li}^u\big[I\{\varepsilon_i \le c_{\kappa_l}\} - \kappa_l\big]K\big((U_i-u)/h\big)\right)\\
&\quad + o_P\Big(\frac{1}{nh}\Big),
\end{aligned}
\tag{A.45}
$$

where $l_1 = p+1, l_2^* = 2p+q+1$, $\mathbf{\Psi}_{f_\varepsilon}(u) = E[(f_\varepsilon^\tau(c_\kappa), c_{f_\varepsilon}\mathbf{X}^\tau)^\tau(\boldsymbol{\xi}^{(2)}(V))^\tau|U = u]\mathbf{C}\boldsymbol{\beta}$. Together with (A.44) and (A.45), using the convexity lemma (Pollard, 1991) and the quadratic approximation lemma (Fan and Gijbels, 1996), we have

$$\sqrt{nh}\left\{\begin{pmatrix}\hat{\alpha}_{01}(u)-\alpha_0(u)-c_{\kappa_1}\\ \vdots \\ \hat{\alpha}_{0q}(u)-\alpha_0(u)-c_{\kappa_q}\\ \hat{\boldsymbol{\alpha}}(u)-\boldsymbol{\alpha}(u)\end{pmatrix} - \frac{h^2\mu_{K_2}}{2}\begin{pmatrix}c\mathbf{1}_{q\times 1}\alpha_0''(u)\\ \boldsymbol{\alpha}''(u)\end{pmatrix} + \frac{h_o^2\mu_{L_2}}{2}\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\mathbf{\Psi}_{f_\varepsilon}(u)\right\} \xrightarrow{\mathcal{L}} N\left(\mathbf{0}_{(q+p)\times 1}, \frac{\vartheta_{K_0}}{f_U(u)}\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\mathbf{G}(u)\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\right), \tag{A.46}$$

where $\mathbf{G}(u) = \sum_{l=1}^q\sum_{s=1}^q(\kappa_l\wedge\kappa_s - \kappa_l\kappa_s)E[(m_l^\tau, \mathbf{X}^\tau)^\tau(m_s^\tau, \mathbf{X}^\tau)|U = u]$. Define $\boldsymbol{\pi}_{q,1} = (\mathbf{1}_q^\tau, \mathbf{0}_p^\tau)^\tau$, using (A.46), recalling that $\hat{\alpha}_0(u) = \frac{1}{q}\sum_{l=1}^q\hat{\alpha}_{0l}(u)$, we have

$$\sqrt{nh}\left(\hat{\alpha}_0(u)-\alpha_0(u)-\frac{1}{q}\sum_{l=1}^q c_{\kappa_l} - \frac{h^2\mu_{K_2}}{2}\alpha_0''(u) - \frac{h_o^2\mu_{L_2}}{2q}\boldsymbol{\pi}_{q,1}^\tau\mathbf{A}_{2,f_\varepsilon(u)}^{-1}\mathbf{\Psi}_{f_\varepsilon}(u)\right) \xrightarrow{\mathcal{L}} N\left(0, \frac{\vartheta_{K_0}}{f_U(u)q^2}\boldsymbol{\pi}_{q,1}^\tau\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\mathbf{G}(u)\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\boldsymbol{\pi}_{q,1}\right). \tag{A.47}$$

Define $\boldsymbol{\pi}_{p,2} = (\mathbf{0}_{p\times q}, \mathbf{I}_p)$, similar to (A.47), we have

$$\sqrt{nh}\left(\hat{\boldsymbol{\alpha}}(u)-\boldsymbol{\alpha}(u)-\frac{h^2\mu_{K_2}}{2}\boldsymbol{\alpha}''(u) - \frac{h_o^2\mu_{L_2}}{2}\boldsymbol{\pi}_{p,2}\mathbf{A}_{2,f_\varepsilon(u)}^{-1}\mathbf{\Psi}_{f_\varepsilon}(u)\right) \xrightarrow{\mathcal{L}} N\left(0, \frac{\vartheta_{K_0}}{f_U(u)}\boldsymbol{\pi}_{p,2}\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\mathbf{G}(u)\mathbf{A}_{2,f_\varepsilon}^{-1}(u)\boldsymbol{\pi}_{p,2}^\tau\right). \tag{A.48}$$

We complete the proof of Theorem 6.

Acknowledgments

The authors thank the Editor, the Associate Editor and a referee for their constructive suggestions that helped them to improve the early manuscript. Yan Zhou's research was supported by the National Science Foundation of China (Grant No. 11701385), National Statistical Research Project (Grant No. 2017LY56), and Doctor start fund of Guangdong Province (Grant No. 2016A030310062). Xia Cui is a College Talent Cultivated by "Thousand-Hundred-Ten" Program of Guangdong

Province and her research is supported by the National Natural Science Foundation of China (Grant No. 11471086), Humans and Social Science Research Team of Guangzhou University (Grant No. 201503XSTD) and the Training Program for Excellent Young College Teachers of Guangdong Province (Grant No.Yq201404). Wangli Xu's research is supported by the National Natural Science Foundation of China (Grant No.11471335), and the MOE Project of Key Research Institute of Humanities and Social Sciences at Universities (16JJD910002).

References

Andrews, D. and Herzberg, A. (1985). *Data. A Collection of Problems for Many Fields for Student and Research Worker*. New York: Springer.

Bravo, F. (2013). Partially linear varying coefficient models with missing at random responses. *Annals of the Institute of Statistical Mathematics* **65**, 721–762. MR3094954

Cai, Z., Fan, J. and Li, R. (2000). Efficient estimation and inferences for varying-coefficient models. *Journal of the American Statistical Association* **95**, 888–902. MR1804446

Cai, Z., Naik, P. A. and Tsai, C.-L. (2000). Denoised least squares estimators: An application to estimating advertising effectiveness. *Statistica Sinica* **10**, 1231–1241. MR1804548

Cai, Z. and Xiao, Z. (2012). Semiparametric quantile regression estimation in dynamic models with partially varying coefficients. *Journal of Econometrics* **167**, 413–425. MR2892085

Carroll, R. J., Fan, J., Gijbels, I. and Wand, M. P. (1997). Generalized partially linear single-index models. *Journal of the American Statistical Association* **92**, 477–489. MR1467842

Carroll, R. J., Ruppert, D., Stefanski, L. A. and Crainiceanu, C. M. (2006). *Nonlinear Measurement Error Models, a Modern Perspective*, 2nd ed. New York: Chapman and Hall. MR2243417

Cui, H., He, X. and Zhu, L. (2002). On regression estimators with de-noised variables. *Statistica Sinica* **12**, 1191–1205. MR1947071

Cui, H. and Hu, T. (2011). On nonlinear regression estimator with denoised variables. *Computational Statistics & Data Analysis* **55**, 1137–1149. MR2736500

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications, Vol. 66*. London: Chapman & Hall. MR1383587

Fan, J. and Zhang, W. (2008). Statistical methods with varying coefficient models. *Statistics and its Interface* **1**, 179–195. MR2425354

Fan, Y. and Zhu, L. (2013). Estimation of general semi-parametric quantile regression. *Journal of Statistical Planning and Inference* **143**, 896–910. MR3011302

Feng, L., Zou, C. and Wang, Z. (2012). Local Walsh-average regression. *Journal of Multivariate Analysis* **106**, 36–48. MR2887678

Fuller, W. A. (1987). *Measurement Error Models*. New York: Wiley. MR0898653

Gu, J. and Liang, Z. (2014). Testing cointegration relationship in a semiparametric varying coefficient model. *Journal of Econometrics* **178**, 57–70. MR3137892

Guo, J., Tang, M., Tian, M. and Zhu, K. (2013). Variable selection in high-dimensional partially linear additive models for composite quantile regression. *Computational Statistics & Data Analysis* **65**, 56–67. MR3064943

Härdle, W., Liang, H. and Gao, J. (2000). *Partially Linear Models*. Heidelberg: Physica-Verlag. MR1787637

Hastie, T. and Tibshirani, R. (1993). Varying-coefficient models (with discussion). *Journal of the Royal Statistical Society, Series B* **55**, 757–796. MR1229881

He, X. and Liang, H. (2000). Quantile regression estimates for a class of linear and partially linear errors-in-variables models. *Statistica Sinica* **10**, 129–140. MR1742104

Heckman, N. E. (1986). Spline smoothing in partly linear models. *Journal of the Royal Statistical Society, Series B* **48**, 244–248. MR0868002

Hu, Y., Gramacy, R. B. and Lian, H. (2013). Bayesian quantile regression for single-index models. *Statistics and Computing* **23**, 437–454. MR3070402

Jiang, R., Zhou, Z.-G., Qian, W.-M. and Chen, Y. (2013). Two step composite quantile regression for single-index models. *Computational Statistics & Data Analysis* **64**, 180–191. MR3061897

Jiang, X., Jiang, J. and Song, X. (2012). Oracle model selection for nonlinear models based on weighted composite quantile regression. *Statistica Sinica* **22**, 1479–1506. MR3027096

Kai, B., Li, R. and Zou, H. (2010). Local composite quantile regression smoothing: An efficient and safe alternative to local polynomial regression. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **72**, 49–69. MR2751243

Kai, B., Li, R. and Zou, H. (2011). New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models. *The Annals of Statistics* **39**, 305–332. MR2797848

Knight, K. (1998). Limiting distributions for L_1 regression estimators under general conditions. *The Annals of Statistics* **26**, 755–770. MR1626024

Koenker, R. (2005). *Quantile Regression*. Cambridge: Cambridge Univ. Pres. MR2268657

Li, G., Feng, S. and Peng, H. (2011). A profile-type smoothed score function for a varying coefficient partially linear model. *Journal of Multivariate Analysis* **102** 372–385. MR2739122

Li, G., Xue, L. and Lian, H. (2011). Semi-varying coefficient models with a diverging number of components. *Journal of Multivariate Analysis* **102**, 1166–1174. MR2805656

Li, Q., Ouyang, D. and Racine, J. S. (2013). Categorical semiparametric varying-coefficient models. *Journal of Applied Econometrics* **28**, 551–579. MR3064528

Li, R. and Liang, H. (2008). Variable selection in semiparametric regression modeling. *The Annals of Statistics* **36**, 261–286. MR2387971

Liang, H. (2008). *Related Topics in Partially Linear Models*. Saarbrucken, Germany: VDM Verlag.

Mack, Y. P. and Silverman, B. W. (1982). Weak and strong uniform consistency of kernel regression estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **61**, 405–415. MR0679685

Pollard, D. (1984). *Convergence of Stochastic Processes*. *Springer Series in Statistics*. New York: Springer. MR0762984

Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory* **7**, 186–199. MR1128411

Robinson, P. M. (1988). Root n-consistent semiparametric regression. *Econometrica* **56**, 931–954. MR0951762

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley. MR0595165

Shang, S., Zou, C. and Wang, Z. (2012). Local Walsh-average regression for semiparametric varying-coefficient models. *Statistics & Probability Letters* **82**, 1815–1822. MR2956621

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. *Monographs on Statistics and Applied Probability* **26**. London: Chapman and Hall. MR0848134

Speckman, P. E. (1988). Kernel smoothing in partial linear models. *Journal of the Royal Statistical Society, Series B* **50**, 413–436. MR0970977

Sun, J., Gai, Y. and Lin, L. (2013). Weighted local linear composite quantile estimation for the case of general error distributions. *Journal of Statistical Planning and Inference* **143**, 1049–1063. MR3029232

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. *Springer Series in Statistics*. New York: Springer. With applications to statistics. MR1385671

Wang, H. and Xia, Y. (2009). Shrinkage estimation of the varying coefficient model. *Journal of the American Statistical Association* **104**, 747–757. MR2541592

Wang, H. J., Stefanski, L. A. and Zhu, Z. (2012). Corrected-loss estimation for quantile regression with covariate measurement errors. *Biometrika* **99**, 405–421. MR2931262

Wang, H. J., Zhu, Z. and Zhou, J. (2009). Quantile regression in partially linear varying coefficient models. *The Annals of Statistics* **37**, 3841–3866. MR2572445

Wang, L., Kai, B. and Li, R. (2009). Local rank inference for varying coefficient models. *Journal of the American Statistical Association* **104**, 1631–1645. MR2597005

Wei, F., Huang, J. and Li, H. (2011). Variable selection and estimation in high-dimensional varying-coefficient models. *Statistica Sinica* **21**, 1515–1540. MR2895107

Wei, Y. and Carroll, R. J. (2009). Quantile regression with measurement error. *Journal of the American Statistical Association* **104**, 1129–1143. MR2562008

Xia, Y., Tong, H., Li, W. K. and Zhu, L.-X. (2002). An adaptive estimation of dimension reduction space. *Journal of the Royal Statistical Society, Series B, Statistical Methodology* **64**, 363–410. MR1924297

Xia, Y., Zhang, W. and Tong, H. (2004). Efficient estimation for semivarying-coefficient models. *Biometrika* **91**, 661–681. MR2090629

Yuan, Y., Zhu, H., Styner, M., Gilmore, J. H. and Marron, J. S. (2013). Varying coefficient model for modeling diffusion tensors along white matter tracts. *Annals of Applied Statistics* **7**, 102–125. MR3086412

Zhang, W., Lee, S.-Y. and Song, X. (2002). Local polynomial fitting in semivarying coefficient model. *Journal of Multivariate Analysis* **82**, 166–188. MR1918619

Zhou, Y. and Liang, H. (2009). Statistical inference for semiparametric varying-coefficient partially linear models with error-prone linear covariates. *The Annals of Statistics* **37**, 427–458. MR2488358

Zhu, H., Li, R. and Kong, L. (2012). Multivariate varying coefficient model for functional responses. *The Annals of Statistics* **40**, 2634–2666. MR3097615

Zou, H. and Yuan, M. (2008). Composite quantile regression and the oracle model selection theory. *The Annals of Statistics* **36**, 1108–1126. MR2418651

J. Zhang
College of Mathematics and Statistics
Institute of Statistical Sciences
Shenzhen University
518060, Shenzhen
China
and
Shenzhen-Hong Kong Joint Research Center for Applied Statistical Sciences
Shenzhen University
518060, Shenzhen
China

Y. Zhou
College of Mathematics and Statistics
Institute of Statistical Sciences
Shenzhen University
518060, Shenzhen
China
E-mail: zhouy1016@szu.edu.cn

X. Cui
School of Economics and Statistics
Guangzhou University
510006, Guangzhou
China

W. Xu
Center for Applied Statistics
School of Statistics
Renmin University of China
Beijing, 100872
China