Quantile Regression for Single-index Varying-coefficient Models with Missing Covariates at Random

Xiaobo Ji, Shuanghua Luo*, Meijuan Liang

Abstract—**This paper studies the single-index varying-coefficient quantile model with missing covariates at random. Firstly, some estimators of index parameters and their corresponding linkage function are given by using the inverse probability weighting method for missing data in two cases including parameter estimation and non-parametric estimation for the single-index varying-coefficient quantile regression model. In particular, the latter case focuses on the study of both known and unknown probability functions. Secondly, the established estimators are proved to be asymptotic normal under some suitable regularity conditions. Finally, the simulation studies are conducted to demonstrate the finite sample performance of the proposed method.**

Index Terms—**quantile regression, inverse probability weighting, missing covariates at random, single-index varying-coefficient model.**

I. INTRODUCTION

AS an important semi-parametric model with index items, single-index varying-coefficient model (SIVCM) is proposed firstly by Xia and Li [1]. This model immediately attracted the attention of many scholars upon its proposal. This is because of its two main advantages: one is that it has the explanatory power like some parametric models, another is that it can avoid the curse of dimensionality.

The SIVCM follows the general form:

$$Y = \eta^T(\alpha^T X)G + \varepsilon, \qquad (1)$$

where Y is the response variable, $X \in R^p$ and $G \in R^d$ are p-dimensional and d-dimensional covariates, respectively. $\eta(\cdot) = (\eta_1(\cdot), \cdots, \eta_d(\cdot))^T$ is a d-dimensional unknown cofficient function vector, and α is the p-dimensional unknown parameter vector. Generally, the first component of G always be taken as 1. And for the sake of identifiability [2], we assume that $\|\alpha\| = 1$ and the first component of α is positive. The $\|\cdot\|$ denotes the Euclidean norm, and ε is a random error independent of (X, G).

Model (1) is so flexible that statistical inference about it has received lots of attentions in literatures. For example, Xue and Wang [3] discussed SIVCM by empirical likelihood method. Huang, et al [4] proposed a procedure for model structure selection in the framework of the SIVCM. Zhao, et al [5] discussed a robust and effective estimation procedure for SIVCM by combining minimum average variance estimation (MAVE) with exponential squared loss. Other work about SIVCM can be seen in [6]-[7].

However, the models mentioned above based on mean regression are not robust against the outliers. Koenker and Bassett [8] proposed quantile regression that can effectively overcome the impact of outliers and non-normal error. Recently, Kuruwita [9] disscussed the variables selection of the single-index quantile regression model with high dimensional covariates. And Xu, et al [10] considered the single-index quantile regression under left truncated data. For more work about single-index varying-coefficient quantile regression (SIVCQR), see [11]-[13].

In addition, missing data is a common issue in social, economic and biomedical studies. To overcome the impact of missing data on estimation results, scholars have proposed some methods such as complete case (CC) analysis, inverse probability weighting (IPW), imputation methods and so on. For model (1) with missing data at random, there have been many researches to concern the estimation for this model by using the inverse probability weighting (IPW) methods. For example, Zhao [15] discussed the estimation of model (1) and used IPW to construct a weighted estimator for the index parameters with missing covariates. Song, et al [16] investigated the robust variable selection for SIVCM and adopted the IPW method to eliminate the potential bias. We also can get robust estimation based on IPW in other models with missing data, such as [17]-[19]. Thus, this paper will adopt IPW to handle the single-index varying-coefficient quantile regression model (SIVCQRM) with missing covariates at random.

In this paper, parameter estimators and non-parametric estimators are proposed for the SIVCQRM with random missing covariates by using several methods including quantile regression (QR) with known selection probability, nonparametric quantile regression (NQR) and parameter quantile regression (PQR) with unknown selection probability, and their asymptotic properties are established under some regularity conditions. Further, the finite sample performance of the proposed method are demonstrated by the simulation studies.

The rest of this paper is organized as follows. In Section II, QR based on IPW utilizes local linear methods, kernel estimation and maximum likelihood estimation respectively under different conditions of SIVCM to obtain the corresponding nonparameter estimators and parameter estimators. The asymptotic properties of established estimators are proved in Section III. The simulation studies are conducted to demonstrate the finite sample performance of the proposed

Manuscript received August 10, 2023; revised April 12, 2024.

This work was supported by the Natural Science Foundation of Shaanxi Province of China (No. 2024JC-YBMS-007).

Xiaobo Ji is a postgraduate student of the School of Science, Xi'an Polytechnic University, Xi'an 710048, China. (email: swdmjxb@163.com).

Shuanghua Luo* is a professor of the School of Science, Xi'an Polytechnic University, Xi'an 710048, China. (email: iwantflyluo@163.com).

Meijua Liang is a postgraduate student of the School of Science, Xi'an Polytechnic University, Xi'an 710048, China. (email: liangmeijuan@163.com).

method in Section IV. The article is briefly discussed and summarized in the Section V.

II. Estimation

Let $V_i = (Y_i, G_i)^T$. Assume that X_i is missing at random, which means the selection probability

$$P(\delta_i = 1|Y_i, X_i, G_i) = P(\delta_i = 1|V_i) = \pi(V_i), \quad (2)$$

where $\delta_i = 0$ if X_i is missing, otherwise $\delta_i = 1$.

Theoretically, when selection probability function $\pi(\cdot)$ is known the QR estimators of $\hat{\alpha}$ can be defined as

$$\hat{\alpha} = \underset{\|\alpha\|=1,\alpha_1>0}{\arg\min} \sum_{i=1}^{n} \frac{\delta_i}{\pi(V_i)} \rho_\tau[(Y - \eta^T(\alpha^T X)G)], \quad (3)$$

if $\eta(\cdot)$ is known, where $\rho_\tau(v) = v(\tau - I_{(v<0)}) = \tau v - vI_{(v<0)}$ is the loss function.

Suppose that $\{x_i, g_i, y_i\}_{i=1}^n$ are independent identically distributed samples from (X, G, Y). For $k = 1, 2, \cdots, d$, $\eta_k(\cdot)$ can be approximated linearly if $\eta(\cdot)$ is unknown. When t in a neighborhood of $\alpha^T x_i$,

$$\begin{aligned}\eta(\alpha^T x_i) &= \eta_k(t) + \eta_k'(t)(\alpha^T x_i - t)\\ &= a_k + b_k(\alpha^T x_i - t),\end{aligned}$$

where $a_k = \eta_k(t)$, $b_k = \eta_k'(t)$. Then the objective function in (3) can be rewritten as

$$\sum_{i=1}^{n} \frac{\delta_i}{\pi(V_i)} \rho_\tau(y_i - \mathbf{a}g_i - \mathbf{b}g_i(\alpha^T x_i - t)) \times J(\frac{\alpha^T x_i - t}{h}), \quad (4)$$

where $\mathbf{a} = (a_1, a_2, \cdots, a_d)$, $\mathbf{b} = (b_1, b_2, \cdots, b_d)$, $J(\cdot)$ is the kernel function and h is the bandwidth. By averaging t, one can get the empirical approximation of (4)

$$\sum_{j=1}^{d} \sum_{i=1}^{n} \frac{\delta_i}{\pi(V_i)} \rho_\tau(y_i - \mathbf{a}g_i - \mathbf{b}g_i\alpha^T(x_i - x_j))\omega_{ij}, \quad (5)$$

where

$$w_{ij} = \frac{J_h(\alpha^T x_i - \alpha^T x_j)}{\sum_{l=1}^{n} J_h(\alpha^T x_l - \alpha^T x_j)},$$

$$J_h(\cdot) = J(\cdot/h)/h.$$

Then it follows form (3) and the above formulations that the quantile regression estimator of SIVCM is defined by

$$\begin{aligned}(\hat{\alpha}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) = \underset{\|\alpha\|=1,\alpha_1>0}{\arg\min} &\sum_{j=1}^{d} \sum_{i=1}^{n} \frac{\delta_i}{\pi(V_i)}\\ &\times \rho_\tau(y_i - \mathbf{a}g_i - \mathbf{b}g_i\alpha^T(x_i - x_j))\\ &\times J(\frac{\alpha^T x_i - \alpha^T x_j}{h}),\end{aligned} \quad (6)$$

when selection probability function $\pi(\cdot)$ is known.

However, the selection probability function $\pi(\cdot)$ is often unknown in many cases. Thus it is necessary to estimate the function $\pi(\cdot)$. We often use nonparametric smoothing estimation approaches to estimate the unknown selection probability function. As a common method, the Nadaraya-Watson estimator of $\pi(v)$ can be defined as

$$\hat{\pi}(v) = \frac{\sum_{i=1}^{n} \delta_i L_h(V_i - v)}{\sum_{i=1}^{n} L_h(V_i - v)}, \quad (7)$$

where $L_h(\cdot) = L(\cdot/h)/h$ is a kernel function, and h is the bandwidth. Hence the NQR estimator with $\hat{\pi}(v)$ is defined as

$$\begin{aligned}(\hat{\alpha}_N, \hat{\mathbf{a}}_N, \hat{\mathbf{b}}_N) = \underset{\|\alpha\|=1,\alpha_1>0}{\arg\min} &\sum_{j=1}^{d} \sum_{i=1}^{n} \frac{\delta_i}{\hat{\pi}(V_i)}\\ &\times \rho_\tau(y_i - \mathbf{a}g_i - \mathbf{b}g_i\alpha^T(x_i - x_j))\\ &\times J(\frac{\alpha^T x_i - \alpha^T x_j}{h}),\end{aligned} \quad (8)$$

where $\hat{\alpha}_N$ is called NQR estimator of α with $\hat{\pi}(v)$.

On the other hand, nonparametric estimation may encounter the curse of dimension when the dimension of V_i is too high. As a result, the parameter method can be used to get the estimator of $\pi(\cdot)$. Supposing that $\pi(v) = \pi(v, \omega)$ for function $\pi(\cdot, \omega)$, where $\pi(\cdot)$ is a known function, and ω is an unknown parameter. Assume

$$\pi(v, \omega) = \frac{\exp(v^T\omega)}{1 + \exp(v^T\omega)}, \quad (9)$$

when $\pi(v, \omega)$ is specified correctly, the estimator $\hat{\omega}$ can be obtained by maximum likelihood estimation (MLE). Then the PQR estimators of SIVCM can be defined as

$$\begin{aligned}(\hat{\alpha}_P, \hat{\mathbf{a}}_P, \hat{\mathbf{b}}_P) = \underset{\|\alpha\|=1,\alpha_1>0}{\arg\min} &\sum_{j=1}^{d} \sum_{i=1}^{n} \frac{\delta_i}{\pi(V_i, \hat{\omega})}\\ &\times \rho_\tau(y_i - \mathbf{a}g_i - \mathbf{b}g_i\alpha^T(x_i - x_j))\\ &\times J(\frac{\alpha^T x_i - \alpha^T x_j}{h}),\end{aligned} \quad (10)$$

where $\hat{\alpha}_P$ is called PQR estimator of α with $\pi(V_i, \hat{\omega})$. The objective function in optimization problems is convex, and the resulting PQR estimators are uniquely defined.

III. Asymptotic properties

The asymptotic properties will be established in this section for the proposed estimators. Above all, some regularity conditions will be introduced in the following.

A1. The kernel $J(\cdot)$ is a symmetric density function with finite support.

A2. In a neighborhood of α_0, the density function of $\alpha^T X$ is positive and uniformly continuous for α. Further the density of ${\alpha_0}^T X$ is continuous and bounded away from 0 and ∞ on its support.

A3. $\eta_0(\cdot)$ is a continuous function which has bounded second derivative.

A4. The model error ε has a symmetric distribution with a positive density $f(\cdot)$.

A5. $A_1(t)$ is non-singular for all $t \in \Omega$ and C_1 is positive definite.

A6. The selection probability function $\pi(v)$ is positive and has a bounded continuous second derivative on the support of (Y, G).

A7. The MLE $\hat{\omega}$ of ω is root-n consistent and satisfying the regularity conditions of asymptotic normality.

Conditions (A1)-(A5) are standard conditions and is commonly used in quantile regression, see Wu[20]. And (A6) is a indispensable condition for analyzing missing data , (A7) is a regular condition for MLE.

Lemma1[21]. Assuming that $A_n(s)$ is convex and can be expressed as $\frac{1}{2}s^T Vs + U_n^T s + C_n + r_n s$, where V is symmetrically positive definite, U_n is randomly bounded, C_n is arbitrary, and $r_n s$ approaches 0 with probability for every s, then the least of $A_n(s)$ is α_n only $o_p(1)$ away from the minimization of $\frac{1}{2}s^T Vs + U_n^T s + C_n$ is $\alpha_n = -V^{-1}U_n$.

Lemma2[22]. Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be independent and identically distributed (i.i.d) random vectors, where the Y_i's are scalar random variables. Suppose that $E|Y_i|^3 < \infty$ and $\sup_x \int |y|^s \varphi(x, y)dy < \infty$, where $f(\cdot, \cdot)$ represents the density of (X, Y). Let $J(\cdot)$ be a bounded positive function with a bounded support, satisfying the Lipschitz condition. Then

$$\sup_x \left| \frac{1}{n} \sum_{i=1}^{n} \{J_h(X_i - x)Y_i - E[J_h(X_i - x)Y]\} \right| = O_p(\ln^{\frac{1}{2}}(1/h) \Big/ \sqrt{nh})),$$

as $n^{2\epsilon-1}h \to \infty$ for $\epsilon < 1 - s^{-1}$.

Let $F(\cdot)$ and $f(\cdot)$ be the cumulative distribution function and density function of model error respectively. Using $f_t(\cdot)$ to represent the marginal density function of $T = \alpha_0^T X$. We select $J(\cdot)$ as the symmetric density function, denoted as $\mu_j = \int t^j J(t)dt$, $v_j = \int t^j J^2(t)dt$. The asymptotic normality of $\hat{\eta} = \hat{\eta}(t; h, \hat{\alpha})$ and $\hat{\alpha}$ are stated in the following theorems.

Theorem 1. Assume $\pi(v)$ is known and conditions (A1)-(A6) hold, if $n \to \infty, h \to 0$ and $nh \to \infty$, then for any interior point t,

$$\sqrt{nh}\{\hat{\eta}(t; h, \hat{\alpha}) - \eta_0(t) - \frac{1}{2}\eta_0''(t)\mu_2 h^2\} \to N(0, \Sigma(t)),$$

where $\Sigma(t) = \frac{v_0\tau(1-\tau)}{f_T(t)} A_1(t)^{-1} A_0(t) A_1(t)^{-1}$, $A_0(t) = E\{G_i G_i^T | T = t\}$, and $A_1(t) = E[f(0|G, T)G_i G_i^T | T = t]$.

Proof Note that

$$\begin{aligned} &\sqrt{nh}\{\hat{\eta}(t; h, \hat{\alpha}) - \eta_0(t)\} \\ &= \sqrt{nh}\{\hat{\eta}(t; h, \hat{\alpha}) - \hat{\eta}(t; h, \alpha_0)\} \\ &+ \sqrt{nh}\{\hat{\eta}(t; h, \alpha_0) - \eta_0(t)\}, \end{aligned}$$

where $\hat{\eta}(\cdot; h, \alpha_0)$ is a local linear estimator of $\eta_0(\cdot)$ when the parameter α_0 is known. $\sqrt{nh}\{\hat{\eta}(t; h, \hat{\alpha}) - \hat{\eta}(t; h, \alpha_0)\}$ can be proved as $o_p(1)$. The details are given below.
For given t,

$$\begin{aligned} (\hat{\eta}(t; h, \hat{\alpha}), \hat{\eta}'(t; h, \hat{\alpha})) &= \arg\min_{(\mathbf{a}, \mathbf{b})} \sum_{i=1}^{n} \frac{\delta_i}{\pi(V_i)} \\ &\times \rho_\tau\{y_i - \mathbf{a}^T g_i - \mathbf{b}^T g_i(\hat{\alpha}^T x_i - t)\} \\ &\times J_h(\hat{\alpha}^T x_i - t), \end{aligned}$$

$$\begin{aligned} (\hat{\eta}(t; h, \alpha_0), \hat{\eta}'(t; h, \alpha_0)) &= \arg\min_{(\mathbf{a}, \mathbf{b})} \sum_{i=1}^{n} \frac{\delta_i}{\pi(V_i)} \\ &\times \rho_\tau\{y_i - \mathbf{a}^T g_i - \mathbf{b}^T g_i(\alpha_0^T x_i - t)\} \\ &\times J_h(\alpha_0^T x_i - t), \end{aligned}$$

In the following, some signs are introduced,

$$\bar{w}^* = \sqrt{nh}\{\hat{\eta}(t; h, \hat{\alpha}) - \eta_0(t), h(\hat{\eta}'(t; h, \hat{\alpha}) - \eta_0'(t))\},$$

$$\bar{w}^{**} = \sqrt{nh}\{\hat{\eta}(t; h, \alpha_0) - \eta_0(t), h(\hat{\eta}'(t; h, \alpha_0) - \eta_0'(t))\},$$

$$y_i^* = y_i - \eta_0(t)^T g_i - \{\eta_0'(t)\}^T g_i(\hat{\alpha}^T x_i - t),$$

$$y_i^{**} = y_i - \eta_0(t)^T g_i - \{\eta_0'(t)\}^T g_i(\alpha_0^T x_i - t),$$

$$J_i^* = J_h(\hat{\alpha}^T x_i - t), J_i^{**} = J_h(\alpha_0^T x_i - t),$$

$${G_i}^* = \{g_i, g_i(\hat{\alpha}^T x_i - t)/h\},$$

$${G_i}^{**} = \{g_i, g_i(\alpha_0^T x_i - t)/h\},$$

and then $\bar{w}^*$ and $\bar{w}^{**}$ minimize

$$\psi_n^*(w) = \sum_{i=1}^{n} [\frac{\delta_i}{\pi(V_i)} \rho_\tau(y_i^* - \frac{w^T G_i^*}{\sqrt{nh}}) - \rho_\tau(y_i^*)] J_i^*$$

and

$$\psi_n^{**}(w) = \sum_{i=1}^{n} [\frac{\delta_i}{\pi(V_i)} \rho_\tau(y_i^{**} - \frac{w^T G_i^{**}}{\sqrt{nh}}) - \rho_\tau(y_i^{**})] J_i^{**}$$

respectively. $\psi_n^*(w)$ and $\psi_n^{**}(w)$ are convex with respect to w, further, they both converge point by point to their conditional expectation, the quadratic approximation is easily deduced. Then by using Lemma 1, we can obtain

$$\psi_n^*(w) = \frac{1}{2} w^T S^* w + W_n^{*T} w + o_p(1),$$

$$\psi_n^{**}(w) = \frac{1}{2} w^T S^{**} w + W_n^{**T} w + o_p(1),$$

where

$$S^* = S^{**} = f_{T_0}(t)\varphi''(0|t) \begin{pmatrix} 1 & 0 \\ 0 & \int_v J(v)v^2 dv, \end{pmatrix},$$

$$W_n^* = -(nh)^{-1/2} \sum_{i=1}^{n} \rho_\tau'(y_i^*) G_i^* J_i^*,$$

$$W_n^{**} = -(nh)^{-1/2} \sum_{i=1}^{n} \rho_\tau'(y_i^{**}) G_i^{**} J_i^{**},$$

and $\varphi''(0|t)$ is the second derivative of

$$\varphi(m|t) = E(\rho_\tau(y - \eta_0(t) + m)|T = t)$$

with respect to m evaluated at $m = 0$.
Assume the first derivatives $\varphi'(m|t)$ and the second derivatives $\varphi''(m|t)$ of $\varphi(m|t)$ with respect to m exist. Then $v \in [-M, M]$, and M is a real number which makes $[-M, M]$ involve the support of $K(\cdot)$.
Due to the [23] and Lemma 1, $\bar{w}^*$ that minimize $\psi_n^*(w)$ can be expressed as

$$\bar{w}^* = -\{S^*\}^{-1} {W_n}^* + o_p(1),$$

w can be got same as

$$\bar{w}^{**} = -\{S^{**}\}^{-1} {W_n}^{**} + o_p(1).$$

Then we have

$$\begin{aligned}\bar{w}^* - \bar{w}^{**} &= -\frac{1}{S^*}(W_n{}^* - W_n{}^{**}) + o_p(1)\\ &= -\frac{1}{S^*}\sum_{i=1}^{n}[(\rho_\tau'(y_i^*)G_i^*J_i^* - \rho_\tau'(y_i^{**})G_i^{**}J_i^{**})]\\ &= -\frac{1}{S^*}\sum_{i=1}^{n}\rho_\tau'(y_i^*)G_i^*J_i^* - \rho_\tau'(y_i^{**})G_i^{**}J_i^{**},\end{aligned}$$

y_i^{**} has the same sign as y_i^* a.s. when

$$\|\hat{\alpha} - \alpha_0\| = O_p(n^{-1/2}).$$

For some $r > 0$,

$$\begin{aligned}&E\{(\bar{w}^* - \bar{w}^{**})(\bar{w}^* - \bar{w}^{**})^T\}\\ &\leq -r\{S^*\}^{-1}h^{-1}(E\{((\rho_\tau'(y_i^*))^2\\ &\times (G_i^*J_i^* - G_i^{**}J_i^{**})(G_i^*J_i^* - G_i^{**}J_i^{**})^T\}\\ &\times (\{S^*\}^{-1})^T\\ &= O(h^{-1}E\{(G_i^*J_i^* - G_i^{**}J_i^{**})\\ &\times ((G_i^*J_i^* - G_i^{**}J_i^{**})^T\})\\ &= O(o(1)) = o(1),\end{aligned}$$

which means $E(\bar{w}^* - \bar{w}^{**}) = o(1)$. Then we can obtain $(\bar{w}^* - \bar{w}^{**}) = E(\bar{w}^* - \bar{w}^{**}) + o_p(1) = o_p(1)$ due to the first two terms. So $(\bar{w}^* - \bar{w}^{**}) = o_p(1)$ and $\sqrt{nh}\{\hat{\eta}(t;h,\hat{\alpha}) - \hat{\eta}(t;h,\alpha_0)\} = o_p(1)$. Then we need to prove that

$$\begin{aligned}&\sqrt{nh}\{\hat{\eta}(t;h,\alpha_0) - \eta_0(t) - \frac{1}{2}\eta_0''(t)\mu_2h^2\}\\ &\to N(0,\Sigma(t)).\end{aligned}$$

The details are given as follows.
Let $\bar{w}^{**} = \sqrt{nh}\{(\hat{\mathbf{a}} - \eta_0(t))^T, h(\hat{\mathbf{b}} - \eta_0'(t))^T\}^T$ and $\bar{w}^{**}$ is the minimizer of the following formulation

$$\begin{aligned}\psi_n(\bar{w}^{**}) = \sum_{i=1}^{n}[&\frac{\delta_i}{\pi(V_i)}\rho_\tau(\varepsilon_i + r_i - \Delta_i)\\ &- \rho_\tau(\varepsilon_i + r_i)]J_i,\end{aligned}$$

where $H_i = (g_i, g_i(\alpha_0^Tx_i - t)/h)^T$, $\Delta_i = \bar{w}^{**}H_i/\sqrt{nh}$, $J_i = J((\alpha_0^Tx_i - t)/h)$, $r_i = \eta_0^T(\alpha_0^Tx_i)g_i - \eta_0^T(t)g_i - (\eta_0'(t))^T(\alpha_0^Tx_i - t)g_i$.
By referring to the identity

$$\begin{aligned}&\rho_\tau(u-v) - \rho_\tau(u)\\ &= -v\varphi_\tau(u) + \int_0^v\{I_{(u\leq s)} - I_{(u\leq 0)}\}ds,\end{aligned}$$

where $\varphi_\tau(u) = \tau - I_{(u\leq 0)}$. Then $\psi_n(\bar{w}^{**})$ can be rewritten as

$$\psi_n(\hat{w}) = \psi_{1n}(\bar{w}^{**}) + \psi_{2n}(\bar{w}^{**}),$$

where

$$\psi_{1n}(\bar{w}^{**}) = \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\frac{\delta_iJ_i}{\pi(V_i)}\Delta_i[I_{(\varepsilon_i\leq -r_i)} - \tau],$$

$$\psi_{2n}(\bar{w}^{**}) = \sum_{i=1}^{n}\frac{\delta_iJ_i}{\pi(V_i)}\int_0^{\Delta_i}[I_{(\varepsilon_i\leq -r_i+s)} - I_{(\varepsilon_i\leq -r_i)}]ds.$$

It is easy to get

$$\psi_{2n}(\bar{w}^{**}) = \frac{f_T(t)}{2}\bar{w}^{**}E(S_n)(\bar{w}^{**})^T + o_p(1),$$

$$S_n = \frac{1}{nh}\sum_{i=1}^{n}f(-r_i|G,T)H_i^TH_iJ_i(t),$$

and

$$\begin{aligned}E(S_n) &= E((f(0|G,T)\begin{bmatrix} GG^T & 0\\ 0 & GG^T\mu_2\end{bmatrix}|T)\\ &= S.\end{aligned}$$

Thus,

$$\psi_{2n}(\bar{w}^{**}) = \frac{1}{2}f_T(t)\bar{w}^{**}S(\bar{w}^{**})^T + o_p(1).$$

It follows from the convexity Lemma [23] that, for any compact set, the quadratic approximation to $\psi_n(\bar{w}^{**})$ holds uniformly for $\bar{w}^{**}$ in any compact set, which generates

$$\begin{aligned}\bar{w}^{**} &= -f_T^{-1}(t)S^{-1}\\ &\times \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}J_iH_i[I_{(\varepsilon_i\leq -r_i)} - \tau] + o_p(1).\end{aligned}$$

At this point, S is a quasi-diagonal matrix.

$$\begin{aligned}\sqrt{nh}(\hat{\eta}(t) - \eta_0(t)) &= -f_T^{-1}(t)A_1(t)^{-1}\\ &\times \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}J_iG_i[I_{(\varepsilon_i\leq -r_i)} - \tau] + o_p(1).\end{aligned}$$

So we can acquire

$$E[\sqrt{nh}(\hat{\eta}(t) - \eta_0(t))] = -\frac{1}{2}\eta_0''(u)\mu_2h^2,$$

$$\begin{aligned}&Var[\sqrt{nh}(\hat{\eta}(t) - \eta_0(t))]\\ &= \frac{v_0\tau(1-\tau)}{f_T(t)}A_1(t)^{-1}A_0(t)^{-1}A_1(t)^{-1}.\end{aligned}$$

The proof has been finished.

Theorem 2. Under the same conditions as in Theorem 1 and assuming that $\pi(v)$ is a smoothing function of v and $\pi(v) \geq \varsigma > 0$, we have

$$\begin{aligned}&\sqrt{nh}\{\hat{\eta}_N(t;h,\hat{\alpha}_N) - \eta_0(t) - \frac{1}{2}\eta_0''(t)\mu_2h^2\}\\ &\to N(0,\Sigma^*(t)),\end{aligned}$$

where

$$\Sigma^*(t) = \frac{v_0\tau(1-\tau)}{f_T(t)}A_1(t)^{-1}A_0^*(t)A_1(t)^{-1},$$

$$\begin{aligned}A_0^*(t) = E\{&\frac{\delta_i}{\pi(V_i)}G_iG_i^T|T=t\}\\ &- E\{\frac{1-\pi(V_i)}{\pi(V_i)}E[G_i^T|V_i]^{\otimes 2}|T=t\},\end{aligned}$$

$$\begin{aligned}&A_1(t)^{-1}A_0^*(t)A_1(t)^{-1}\\ &\leq A_1(t)^{-1}E[G_iG_i^T|T=t]A_1(t)^{-1}\\ &= A_1(t)^{-1}A_0(t)A_1(t)^{-1}.\end{aligned}$$

Proof Let

$$\hat{w}_N = \sqrt{nh}\{(\hat{\mathbf{a}}_N - \eta_0(t))^T, h(\hat{\mathbf{b}}_N - \eta_0'(t))^T\}^T.$$

Similarly to the proof of Theorem 1, we have

$$\begin{aligned}\psi_n^*(\hat{\pi}(V_i), w_N) &= \frac{1}{\sqrt{nh}}\sum_{i=1}^{n}\frac{\delta_iJ_i}{\hat{\pi}(V_i)}\\ &\times \{w_N^TH_i[I_{(\varepsilon_i< -r_i)} - \tau]\\ &+ \int_0^{\Delta_i^*}[I_{(\varepsilon_i\leq -r_i+s)} - I_{(\varepsilon_i\leq -r_i)}]ds\}\\ &= \psi_{1n}^*w_N + \psi_{2n}^*(w_N),\end{aligned}$$

where

$$\psi_{1n}^* w_N = \frac{1}{\sqrt{nh}} \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i)} w_N^T H_i [I_{(\varepsilon_i < -r_i)} - \tau],$$

$$\psi_{2n}^*(w_N) = \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i)} \int_0^{\Delta_i^*} [I_{(\varepsilon_i \le -r_i + s)} - I_{(\varepsilon_i \le -r_i)}] ds.$$

Let

$$B_n^*(w_N) = \sum_{i=1}^n \frac{\delta_i J_i (\pi(V_i) - \hat{\pi}(V_i))}{\hat{\pi}(V_i)\pi(V_i)} \times \int_0^{\Delta_i^*} [I_{(\varepsilon_i \le -r_i + s)} - I_{(\varepsilon_i \le -r_i)}] ds,$$

and $\psi_{2n}^*(w_N) = \psi_{2n}(w_N) + B_n^*(w_N)$. It is easy to get that

$$\frac{1}{\sqrt{nh}} \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i)} \int_0^{\Delta_i^*} [I_{(\varepsilon_i \le -r_i + s)} - I_{(\varepsilon_i \le -r_i)}] ds = O_p(1).$$

Because of the fact that

$$\sup_v |\hat{\pi}(v) - \pi(v)| = o(1),$$

it following the above formulations, $B_n^*(w_N) = o_p(1)$. Similarly to the proof of Theorem 1, we can prove that

$$\sqrt{nh}(\hat{\mathbf{a}}_N - \eta_0(t)) = -f_T^{-1}(t) A_1(t)^{-1} \times \frac{1}{\sqrt{nh}} \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i)} G_i \times [I_{(\varepsilon_i \le -r_i)} - \tau] + O_p(1).$$

Let

$$\tilde{\psi}_{1n}^*(t) = \frac{1}{\sqrt{nh}} \sum_{i=1}^n \frac{\delta_i J_i}{\hat{\pi}(V_i)} G_i [I_{(\varepsilon_i \le -r_i)} - \tau].$$

By the proof of Theorem 2 in [18], we can obtain

$$\begin{aligned} \tilde{\psi}_{1n}^*(t) &= \frac{1}{\sqrt{nh}} \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i)} G_i [I_{(\varepsilon_i \le -r_i)} - \tau] \\ &+ \frac{1}{\sqrt{nh}} \sum_{i=1}^n \frac{\delta_i J_i - \pi(V_i)}{\pi(V_i)} \\ &\times E[(I_{(\varepsilon_i \le -r_i)} - \tau) G_i | V_i] + o_p(h^2) \\ &= \tilde{\psi}_{1n,1}^*(t) + \tilde{\psi}_{1n,2}^*(t) + o_p(h^2), \end{aligned}$$

where $E(\tilde{\psi}_{1n,1}^*(t)) = 0$ and $E(\tilde{\psi}_{1n,2}^*(t)) = 0$. In addition, by completing some calculations, we can obtain

$$\begin{aligned} \mathrm{var}(\tilde{\psi}_{1n}^*(t)) &= \frac{v_0}{f_T(t)} \{ E(\frac{\delta_i}{\pi(V_i)} [I_{(\varepsilon_i \le -r_i)} - \tau]^2 \\ &\times G_i G_i^T | T = t) - E[\frac{1 - \pi(V_i)}{\pi(V_i)} \\ &\times E(G_i^T [I_{(\varepsilon_i \le -r_i)} - \tau] | V_i)^{\otimes 2} | T = t] \} \\ &+ o(1). \end{aligned}$$

Then $\tilde{\psi}_{1n,2}^* \xrightarrow{d} N(0, \frac{v_0 \tau(1-\tau)}{f_T(t)} A_0^*(t))$. Following [17], we can get

$$\begin{aligned} &Var(\tilde{\psi}_{1n,2}^* - \psi_{1n,2}^* | X, G) \\ &\le \frac{q^2}{nh} \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i)} G_i G_i^T \max_k \{ F(c_k + |r_i|) - F(c_k) \} \\ &= o_p(1). \end{aligned}$$

Based on above results, it follows that $\tilde{\psi}_{1n}^*(t) \xrightarrow{d} N(0, \frac{v_0 \tau(1-\tau)}{f_T(t)} A_0^*(t))$. By Slutsky's theorem,

$$\tilde{\psi}_{1n}^*(t) - E[\tilde{\psi}_{1n}^*(t)] \xrightarrow{d} N(0, \frac{v_0 \tau(1-\tau)}{f_T(t)} A_0^*(t)).$$

By Lemma 2, we can get

$$\frac{1}{nh} \sum_{i=1}^n \delta_i J_i [I_{(\varepsilon_i \le -r_i)} - \tau] G_i \xrightarrow{p} E[\frac{1}{nh} \sum_{i=1}^n \delta_i J_i [I_{(\varepsilon_i \le -r_i)} - \tau] G_i = O\left(h^2\right).$$

Since $\frac{1}{\hat{\pi}(V_i)} - \frac{1}{\pi(V_i)} = o_p(1)$, then

$$\begin{aligned} \frac{1}{\sqrt{nh}} \tilde{\psi}_{1n}^*(t) &= \frac{1}{\sqrt{nh}} \sum_{i=1}^n \frac{\delta_i J_i}{\hat{\pi}(V_i)} \\ &\times [I_{(\varepsilon_i \le -r_i)} - \tau] G_i \\ &+ \frac{1}{\sqrt{nh}} \sum_{i=1}^n \delta_i J_i [\frac{1}{\hat{\pi}(V_i)} - \frac{1}{\pi(V_i)}] \\ &\times [I_{(\varepsilon_i \le -r_i)} - \tau] G_i \\ &= \frac{1}{nh} \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i)} \eta_i G_i + o_p(h^2). \end{aligned}$$

Thus, we can show that

$$\frac{1}{\sqrt{nh}} E[\tilde{\psi}_{1n}^*(t)] = \frac{1}{\sqrt{nh}} E[\sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i)} \eta_i(u) G_i] + o(h^2).$$

Following above proof and Theorem 1, Theorem 2 is proved.

Theorem 3. Assuming $\pi(v) \ge \varsigma > 0$ is with an unknown parameter ω, based on (A1)-(A7), we have

$$\sqrt{nh} \{ \hat{\eta}_P(t; h, \hat{\alpha}_P) - \eta_0(t) - \frac{1}{2} \eta_0''(t) \mu_2 h^2 \} \to N(0, \Sigma^{**}(t)),$$

where

$$\begin{aligned} &A_1(t)^{-1} A_0^*(t) A_1(t)^{-1} \\ &\le A_1(t)^{-1} E[G_i G_i^T | T = t] A_1(t)^{-1} \\ &= A_1(t)^{-1} A_0(t) A_1(t)^{-1}, \end{aligned}$$

$$\Sigma^{**}(t) = \frac{v_0 \tau(1-\tau)}{f_T(t)} A_1(t)^{-1} A_0^{**}(t) A_1(t)^{-1},$$

$$A_0^{**}(t) = E\{ \frac{\delta_i}{\pi(V_i, \omega)} G_i G_i^T | T = t \} - \Omega_t^T \Lambda_t^{-1} \Omega_t,$$

$$\Omega_t = E[(1 - \pi(V_i, \omega) G_i^T V_i | T = t],$$

$$\Lambda_t = E[V_i^T V_i \pi(V_i, \omega)(1 - \pi(V_i, \omega) | T = t].$$

Proof Let

$$\hat{w}_P = \sqrt{nh} \{ (\hat{\mathbf{a}}_P - \eta_0(t))^T, h(\hat{\mathbf{b}}_P - \eta_0'(t))^T \}^T.$$

Then

$$\begin{aligned} \psi_n^{**}(\pi(V_i, \hat{\omega}), \hat{w}_P) &= \frac{1}{\sqrt{nh}} \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i, \hat{\omega})} w_P^T H_i \\ &\times [I_{(\varepsilon_i \le -r_i)} - \tau] \\ &+ \sum_{i=1}^n \frac{\delta_i J_i}{\pi(V_i, \hat{\omega})} \int_0^{\Delta_i^{**}} \\ &\times [I_{(\varepsilon_i \le -r_i + s)} - I_{(\varepsilon_i \le -r_i)}] ds \\ &= \psi_{1n}^{**} w_P + \psi_{2n}^{**}(w_P), \end{aligned}$$

where

$$\psi_{1n}^{**} = \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \frac{\delta_i J_i}{\pi(V_i, \hat{\omega})} H_i G_i [I_{(\varepsilon_i \leq -r_i)} - \tau],$$

$$\psi_{2n}^{**}(w_P) = \sum_{i=1}^{n} \frac{\delta_i J_i}{\pi(V_i, \hat{\omega})} \int_0^{\Delta_i^{**}} [I_{(\varepsilon_i \leq -r_i + s)} - I_{(\varepsilon_i \leq -r_i)}] ds \}.$$

Let

$$B_n^{**}(w_P) = \sum_{i=1}^{n} \frac{\delta_i J_i (\pi(V_i, \omega) - \pi(V_i, \hat{\omega}))}{\pi(V_i, \hat{\omega}) \pi(V_i, \omega)} \times \int_0^{\Delta_i^{**}} [I_{(\varepsilon_i \leq -r_i + s)} - I_{(\varepsilon_i \leq -r_i)}] ds.$$

Then $\psi_{2n}^{**}(w_P) = \psi_{2n}(w_P) + B_n^{**}(w_P)$. Based on MLE theory and the proof of Theorem 2, we can prove $B_n^{**}(w_P) = o_p(1)$. Similarly we get that

$$\sqrt{nh}(\mathbf{\hat{a}}_P - \eta_0(t)) = -f_T^{-1}(t) A_1(t)^{-1} \frac{1}{\sqrt{nh}} \times \sum_{i=1}^{n} \frac{\delta_i J_i}{\pi(V_i, \hat{\omega})} G_i [I_{(\varepsilon_i \leq -r_i)} - \tau)] + O_p(1).$$

Denote $\pi'(V_i, \omega) = grad(V_i, \omega)$, then

$$\pi(V_i, \hat{\omega}) - \pi(V_i, \omega) = \pi(V_i, \omega)(1 - \pi(V_i, \omega)) \times V_i(\hat{\omega} - \omega) + o_p(n^{-1/2}),$$

$$\sqrt{n}(\hat{\omega} - \omega) = E[\pi'(V_i, \omega)^{\otimes 2}]^{-1} \times \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \pi'(V_i, \omega)^T (\delta_i - \pi(V_i, \omega)).$$

Let

$$\tilde{\psi}_{1n}^{**}(w_P) = \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \frac{\delta_i J_i}{\pi(V_i, \hat{\omega})} G_i [I_{(\varepsilon_i \leq -r_i)} - \tau)].$$

According to the above functions, we have

$$\begin{aligned}
\tilde{\psi}_{1n}^{**}(w_P) &= \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \frac{\delta_i J_i}{\pi(V_i, \omega)} G_i [I_{(\varepsilon_i \leq -r_i)} - \tau)] \\
&\quad - \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \frac{\delta_i J_i (\pi(V_i, \hat{\omega}) - \pi(V_i, \omega))}{\pi(V_i, \omega) \pi(V_i, \hat{\omega})} G_i \\
&\quad \times [I_{(\varepsilon_i \leq -r_i)} - \tau)] \\
&= \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \frac{\delta_i J_i}{\pi(V_i, \omega)} G_i [I_{(\varepsilon_i \leq -r_i)} - \tau)] \\
&\quad - \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \frac{\delta_i J_i (\pi(V_i, \hat{\omega}) - \pi(V_i, \omega))}{\pi^2(V_i, \omega)} \\
&\quad \times G_i [I_{(\varepsilon_i \leq -r_i)} - \tau)] + o_p(1) \\
&= \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \frac{\delta_i J_i}{\pi(V_i, \omega)} G_i [I_{(\varepsilon_i \leq -r_i)} - \tau)] \\
&\quad - \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \frac{\delta_i J_i (1 - \pi(V_i, \omega))}{\pi(V_i, \omega)} G_i \\
&\quad \times [I_{(\varepsilon_i \leq -r_i)} - \tau)] V_i(\hat{\omega} - \omega) + o_p(1) \\
&= \tilde{\psi}_{1n,1}^{**}(w_P) - \tilde{\psi}_{1n,2}^{**}(w_P) + o_p(1).
\end{aligned}$$

Note that $\pi'(v, \omega) = 1 - \pi(v, \omega)$. Then we have

$$\sqrt{n}(\hat{\omega} - \omega) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \pi'(V_i, \omega)^T (\delta_i - \pi(V_i, \omega)) \Lambda_t^{-2}.$$

Since

$$\operatorname{var}(\tilde{\psi}_{1n,1}^{**}(w_P)) = \frac{v_0 \tau(1 - \tau)}{f_T(t)} \times E(\frac{\delta_i}{\pi(V_i, \omega)} G_i G_i^T | T = t),$$

$$\operatorname{var}(\tilde{\psi}_{1n,2}^{**}(w_P)) = \frac{v_0 \tau(1 - \tau)}{f_T(t)} \times \Omega_t^T E[V_i^T V_i \pi(V_i, \omega) \times (1 - \pi(V_i, \omega))] + o(1),$$

$$\operatorname{cov}(\tilde{\psi}_{1n,1}^{**}(w_P), \tilde{\psi}_{1n,2}^{**}(w_P)) = \frac{2 v_0 \tau(1 - \tau)}{f_T(t)} \times \Omega_t^T E[V_i^T V_i \pi(V_i, \omega)(1 - \pi(V_i, \omega))] + o(1),$$

then

$$\begin{aligned}
\operatorname{var}(\tilde{\psi}_{1n}^{**}(w_P)) &= \frac{v_0 \tau(1 - \tau)}{f_T(t)} E(\frac{\delta_i}{\pi(V_i, \omega)} G_i G_i^T | T = t) \\
&\quad - \frac{v_0 \tau(1 - \tau)}{f_T(t)} \Omega_t^T E[V_i^T V_i \pi(V_i, \omega) \\
&\quad \times (1 - \pi(V_i, \omega))] + o(1).
\end{aligned}$$

Similar to the proof of Theorem 2, Theorem 3 follows.

Theorem 4. Assuming conditions (A1)-(A6) hold, if $n \to \infty, h \to 0$ and $nh \to \infty$, then

$$\sqrt{n}(\tilde{\alpha} - \alpha_0) \to N(0, \tau(1 - \tau) C_1^{-1} C_0 C_1^{-1}),$$

where $\tilde{\alpha}$ can be $\hat{\alpha}, \hat{\alpha}_N$ and $\hat{\alpha}_P$,

$$C_0 = E[(X - E(X|\alpha_0^T X)) \eta_0'(\alpha_0^T X)^T \times G_i G_i^T \eta_0'(\alpha_0^T X)(X - E(X|\alpha_0^T X))^T],$$

$$C_1 = E[f(0|X, G)(X - E(X|\alpha_0^T X)) \eta_0'(\alpha_0^T X)^T \times G_i G_i^T \eta_0'(\alpha_0^T X)(X - E(X|\alpha_0^T X))^T].$$

Proof Let $\hat{\gamma} = \sqrt{n}(\tilde{\alpha} - \alpha_0)$. Then $\hat{\gamma}$ is the minimizer of the following criterion

$$Q_n(\gamma) = \sum_{j=1}^{d} \sum_{i=1}^{n} [\frac{\delta_i}{\pi(V_i)} [\rho_\tau(\varepsilon_i - s_i - \frac{1}{\sqrt{n}} \gamma^T x_{ij} \hat{b}_j^T g_i) - \rho_\tau(\varepsilon_i - s_i)] \omega_{ij},$$

where $s_i = -\eta_0^T(\alpha_0^T x_i) g_i + \hat{a}_j^T g_i + \hat{b}_j^T g_i \alpha_0^T x_{ij}$, $x_{ij} = x_i - x_j$.

Thus,

$$\begin{aligned}
Q_n(\gamma) &= E(Q_n(\gamma)) - \frac{1}{\sqrt{n}} \sum_{j=1}^{d} \sum_{i=1}^{n} \gamma^T \frac{\delta_i}{\pi(V_i)} \\
&\quad \times [\omega_{ij} \rho_\tau'(\varepsilon_i - s_i) x_{ij} \hat{b}_j^T g_i \\
&\quad - \omega_{ij} E[\rho_\tau'(\varepsilon_i - s_i)] x_{ij} \hat{b}_j^T g_i] + o_p(1),
\end{aligned}$$

and

$$E[(Q_n(\gamma))] = \sum_{j=1}^{d}\sum_{i=1}^{n}[E\frac{\delta_i}{\pi(V_i)}[\rho_\tau(\varepsilon_i - s_i - \frac{1}{\sqrt{n}}\gamma^T x_{ij}\hat{b}_j^T g_i)] - E[\rho_\tau(\varepsilon_i - s_i]\omega_{ij}$$

$$= \sum_{j=1}^{d}\sum_{i=1}^{n}[E\rho_\tau(y_i - \hat{\eta}(\hat{\gamma}^T x_i|\gamma_{\mathbf{0}}^{\mathbf{T}}\mathbf{x_i} = \gamma_0^T x_i)g_i + \frac{1}{\sqrt{n}}(\hat{\gamma}-\gamma)x_{ij}\hat{b}_j^T g_i)\omega_{ij}$$

$$- \sum_{j=1}^{d}\sum_{i=1}^{n}E\rho_\tau(y_i - \hat{\eta}(\gamma^T x_i|\gamma_{\mathbf{0}}^{\mathbf{T}}\mathbf{x_i} = \gamma_0^T x_i)g_i + \frac{1}{\sqrt{n}}(\hat{\gamma}^T x_{ij}\hat{b}_j^T g_i)\omega_{ij}]$$

$$= -\frac{1}{\sqrt{n}}\gamma^T\sum_{j=1}^{d}\sum_{i=1}^{n}\rho'_\tau(\varepsilon_i - s_i)x_{ij}\hat{b}_j^T g_i\omega_{ij} + \frac{1}{2n}\gamma^T \sum_{j=1}^{d}\sum_{i=1}^{n}2f(0|X,G)x_{ij}\hat{b}_j^T g_i g_i^T \times \hat{b}_j^T x_{ij}^T\omega_{ij}\gamma + o_p(1).$$

As a result, we can acquire

$$Q_n(\gamma) = -\frac{1}{\sqrt{n}}\gamma^T\sum_{j=1}^{d}\sum_{i=1}^{n}\rho'_\tau(\varepsilon_i - s_i)x_{ij}\hat{b}_j^T g_i w_{ij} + \frac{1}{2n}\gamma^T\sum_{j=1}^{d}\sum_{i=1}^{n}2f(0|X,G)x_{ij}\hat{b}_j^T g_i g_i^T \times \hat{b}_j x_{ij}^T w_{ij}\gamma + o_p(1).$$

In the root-n consistency assumption, $\rho'_\tau(\varepsilon_i - s_i)$ has similar asymptotic distribution of $\rho'_\tau(\varepsilon_i)$ given $\hat{a}_j$ and $\hat{b}_j$. Thus, the theorem can be proved.

IV. Numerical studies

In this section, several simulation examples are given to assess the performance of the proposed methods.

In numerical studies, we use the kernel function $J(x) = 0.75(1-x^2)I_{(|x|\leq 1)}$, and it follows from the cross validation method that the optimal bandwidth h_{opt} is selected.

We conduct a small simulation study with $n = 100$ and the data is generated from the following model

$$Y = (X^T\alpha_0)G + (\varepsilon - E_\tau(\varepsilon)),$$

where ε is the model error and $E_\tau(\varepsilon)$ is the τ th quantile of ε, $\alpha_0 = (\alpha_1, \alpha_2, \alpha_3)^T = (2/3, 1/2, 1/3)^T$. The covariate vector $X = (X_1, X_2, X_3)^T$ is a three-dimensionals standard normal variable. The correlation between X_i and X_j is $\frac{1}{2}^{|i-j|}$, $i, j = 1, 2, 3$. X_1 is the missing value and another covariate vector G is generated from a standard normal distribution. In the following simulations, we considered three error distributions: $N(0, 1)$, $t(3)$ and $\chi^2(2)$. All simulations are performed with 500 replicates based on the following selection probability function:

$$P(\delta_i = 1|X_i, G_i) = \frac{\exp(\omega_0+\omega_1 X_2+\omega_2 X_3+\omega_3 G)}{1+\exp(\omega_0+\omega_1 X_2+\omega_2 X_3+\omega_3 G)},$$

where $\omega = (\omega_0, \omega_1, \omega_2, \omega_3) = (-1, 0.3, 0.8, 0.1)$. The average missing rates are approximately 31% when the quantile point are set as 0.25, 0.5 and 0.75.

There are four different estimation methods for the above cases: least square method (LS), quantile regression (QR), quantile regression method under nonparameter estimation (NQR) and parameter estimation (PQR). In these cases, standard deviation (SD) and the mean square error (MSE) of parameter vectors are calculated and simulation results are given in the following tables. Using the same selection probability function and the same sample size, and we can observe that SD of QR, PQR and NQR which are in different quantile points are mostly lower than those of LS. The estimation effect is better under different distributions. The MSE of LS is slightly better than QR method only under

TABLE I
Estimators of SD and MSE on $\tau = 0.25$

Dist	Methods	τ	MEAN			SD			MSE		
			α_1	α_2	α_3	α_1	α_2	α_3	α_1	α_2	α_3
N(0,1)	LS	0.5	0.669	0.505	0.334	0.090	0.091	0.087	0.008	0.008	0.008
	QR	0.25	0.676	0.502	0.333	0.132	0.133	0.120	0.018	0.018	0.014
	PQR	0.25	0.667	0.503	0.330	0.035	0.034	0.030	0.001	0.001	0.001
	NQR	0.25	0.668	0.499	0.333	0.038	0.035	0.037	0.002	0.002	0.002
t(3)	LS	0.5	0.672	0.496	0.328	0.158	0.162	0.145	0.025	0.026	0.021
	QR	0.25	0.678	0.499	0.324	0.155	0.152	0.144	0.026	0.023	0.021
	PQR	0.25	0.665	0.504	0.333	0.040	0.040	0.034	0.002	0.002	0.001
	NQR	0.25	0.665	0.485	0.345	0.307	0.312	0.367	0.094	0.097	0.135
$\chi^2(2)$	LS	0.5	0.672	0.507	0.333	0.166	0.170	0.172	0.028	0.030	0.030
	QR	0.25	0.667	0.497	0.339	0.215	0.232	0.212	0.046	0.054	0.045
	PQR	0.25	0.666	0.502	0.331	0.066	0.060	0.050	0.004	0.004	0.003
	NQR	0.25	0.653	0.490	0.357	0.370	0.369	0.370	0.006	0.006	0.006

TABLE II
Estimators of SD and MSE on $\tau = 0.5$

Dist	Methods	τ	MEAN			SD			MSE		
			α_1	α_2	α_3	α_1	α_2	α_3	α_1	α_2	α_3
N(0,1)	LS	0.5	0.669	0.505	0.334	0.090	0.091	0.087	0.008	0.008	0.007
	QR	0.5	0.669	0.503	0.335	0.104	0.107	0.105	0.011	0.012	0.011
	PQR	0.5	0.667	0.502	0.333	0.029	0.029	0.026	0.001	0.001	0.001
	NQR	0.5	0.668	0.499	0.333	0.041	0.040	0.042	0.002	0.002	0.002
t(3)	LS	0.5	0.672	0.496	0.328	0.158	0.162	0.145	0.025	0.026	0.021
	QR	0.5	0.678	0.496	0.326	0.135	0.135	0.120	0.018	0.018	0.014
	PQR	0.5	0.666	0.502	0.333	0.033	0.035	0.028	0.001	0.001	0.001
	NQR	0.5	0.679	0.489	0.313	0.229	0.231	0.240	0.052	0.053	0.058
$\chi^2(2)$	LS	0.5	0.672	0.507	0.333	0.166	0.170	0.172	0.028	0.029	0.030
	QR	0.5	0.672	0.502	0.338	0.189	0.206	0.192	0.036	0.042	0.037
	PQR	0.5	0.664	0.504	0.332	0.059	0.054	0.046	0.004	0.003	0.002
	NQR	0.5	0.669	0.498	0.334	0.066	0.067	0.066	0.005	0.005	0.004

TABLE III
ESTIMATORS OF SD AND MSE ON $\tau = 0.75$

Dist	Methods	τ	MEAN			SD			MSE		
			α_1	α_2	α_3	α_1	α_2	α_3	α_1	α_2	α_3
N(0,1)	LS	0.5	0.669	0.505	0.334	0.090	0.091	0.087	0.008	0.008	0.008
	QR	0.75	0.668	0.506	0.330	0.117	0.112	0.116	0.014	0.013	0.014
	PQR	0.75	0.667	0.500	0.336	0.037	0.033	0.032	0.001	0.001	0.001
	NQR	0.75	0.665	0.501	0.331	0.049	0.046	0.043	0.002	0.002	0.0018
t(3)	LS	0.5	0.672	0.496	0.328	0.158	0.162	0.145	0.025	0.026	0.021
	QR	0.75	0.670	0.500	0.331	0.156	0.159	0.151	0.024	0.025	0.023
	PQR	0.75	0.667	0.503	0.332	0.040	0.040	0.034	0.002	0.002	0.001
	NQR	0.75	0.665	0.504	0.332	0.051	0.052	0.053	0.003	0.003	0.003
$\chi^2(2)$	LS	0.5	0.672	0.507	0.333	0.166	0.170	0.172	0.028	0.0289	0.030
	QR	0.75	0.666	0.508	0.342	0.226	0.229	0.224	0.051	0.052	0.050
	PQR	0.75	0.668	0.502	0.331	0.068	0.061	0.053	0.005	0.004	0.003
	NQR	0.75	0.671	0.496	0.336	0.072	0.073	0.074	0.005	0.005	0.006

standard normal distribution from above tables. However, the outcomes of QR and PQR are superior to those of the LS in other distributions. Hence, a more robust estimation is provided by the quantile regression in most scenarios, and parameter estimation is better than non-parametric estimation in small sample experiments.

V. CONCLUSION

This paper considers quantile regression estimation of single-index varying-coefficient model with covariates missing at random. The IPW method is used to handle missing covariates. Using different estimation methods to estimate selection probabilities. Numerical simulation results show that methods can achieve good results under different error distributions. And the properties of the large sample estimator and linkage functions are proved.

REFERENCES

[1] Y. C. Xia and W. K. Li. "On single-index coefficient regression models," *Journal of the American Statistical Association*, vol. 94, no. 448, pp. 1275-1285, 1999.

[2] W. Lin and K. B. Kulasekera. "Identifiability of single-index models and additive-index models," *Biometrika*, vol. 94, no. 2, pp. 496-501, 2007.

[3] L. G. Xue and Q. H. Wang. "Empirical likelihood for single-index varying coefficient models," *Bernoulli*, vol. 18, no. 3, pp. 836-856, 2012.

[4] Z. Huang, Z. Pang and B. Lin. "Model structure selection in single-index-coefficient regression models," *Journal of Multivariate Analysis*, vol. 125, pp. 159-175, 2014.

[5] Y. Zhao, L. L. Yue and G. R. Li. "Robust MAVE for single -index varying-coefficient models," *Journal of the Korean Statistical Society*, vol. 51, no. 4, pp. 1302-1325, 2022.

[6] Y. Zhao, L. G. Xue and S. Feng. "Estimation for a partially linear single-index varying-coefficient model," *Communications in Statistics-Simulation and Computation*, vol. 51, no. 4, pp. 1-19, 2019.

[7] Y. Kim. "Efficient estimation and variable selection for partially linear single-index coefficient regression models," *Communications for Statistical Applications and Methods*, vol. 26, no. 1, pp. 69-78, 2019.

[8] R. Koenker and J. G. Bassett. "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33-50, 1978.

[9] C. N. Kuruwita. "Variable selection in the single-index quantile regression model with high dimensional covariates," *Communications in Statistics-Simulation and Computation*, vol. 52, no. 3, pp. 1119-1131, 2023.

[10] H. X. Xu, G. L. Fan and J. C. Li. "Single-index quantile regression with left truncated data," *Journal of Systems Science and Complexity*, vol. 35, no. 5, pp. 1963-1987, 2022.

[11] R. Jiang and W. M. Qian. "Quantile regression for single-index coefficient regression models," *Statistics and Probability Letters*, vol. 110, pp. 305-317, 2016.

[12] J. Yang and H. Yang. " Quantile regression and variable selection for single-index varying-coefficient models," *Communications in Statistics-Simulation and Computation*, vol. 46, no. 6, pp. 4637-4653, 2017.

[13] W. H. Zhao, H. Liang and H. Lian. " Quantile regression for the single-index coefficient model," *Bernoulli: Official Journal of the Bernoulli Society for Mathematical Statistics and Probability*, vol. 23, no. 3, pp. 1997-2027, 2017.

[14] Y. Y. Zhou, G. L. Fan and R. Q. Zhang. "Quantile regression and variable selection for partially linear single-index models with missing censoring indicators," *Journal of Statistical Planning and Inference*, vol. 204, no. 1, pp. 80-95, 2020.

[15] Y. Zhao, L. G. Xue, J. H. Zhang and J. F. Liu. "Single-index varying-coefficient models with missing covariates at random," *Communications in Statistics-Simulation and Computation*, vol. 51, no. 12, pp. 7351-7365, 2022.

[16] Y. Q. Song, Y. Q. Liu and H. Su. "Robust variable selection for single-index varying coefficient model with missing data in covariates," *Mathematics*, vol. 10, no. 12, pp. 2003, 2022.

[17] L. J. Tang and Z. G. Zhou. "Weighted local linear CQR for varying coefficient models with missing covariates," *Test*, vol. 24, no. 3, pp. 583-604, 2015.

[18] S. H. Luo, C. Y. Zhang and M. H. Wang. "Composite quantile regression for varying coefficient models with response data missing at random," *Symmetry*, vol. 11, no. 9, pp. 1065, 2019.

[19] H. Y. Liang, B. H. Wang and Y. Shen. "Quantile regression of partially linear single-index model with missing observations," *Statistics*, vol. 55, no. 1, pp. 1-17, 2021.

[20] T. Wu, K. Yu and Y. Yan. "Single-index quantile regression, "*Journal of Multivariate Analysis*, vol. 101, no. 7, pp. 1607-1621, 2010.

[21] Y. X. Liu, R. X. Rui and M. Z. Tian. "A novel composite quantile regression estimation for the partial linear variable cofficient models," *Acta Mathematicae Applicatae Sinica*, vol. 44, no. 2, pp. 159-174, 2021.

[22] J. Q. Fan and W. Y. Zhang. "Statistical estimation in varying coefficient models," *The Annals of Statistics*, vol. 27, no. 5, pp. 1491-1518, 1999.

[23] P. David. "Asymptotics for least absolute deviation regression estimators," *Econometric Theory*, vol. 7, no. 2, pp. 186-199, 1991.