Taylor & Francis
Taylor & Francis Group

Journal of Business & Economic Statistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ubes20

No-Crossing Single-Index Quantile Regression Curve Estimation

Rong Jiang & Keming Yu

To cite this article: Rong Jiang & Keming Yu (2023) No-Crossing Single-Index Quantile Regression Curve Estimation, Journal of Business & Economic Statistics, 41:2, 309-320, DOI: 10.1080/07350015.2021.2013245

To link to this article: https://doi.org/10.1080/07350015.2021.2013245

JOURNAL OF BUSINESS & ECONOMIC STATISTICS
2023, VOL. 41, NO. 2, 309–320
https://doi.org/10.1080/07350015.2021.2013245

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS   Check for updates

No-Crossing Single-Index Quantile Regression Curve Estimation

Rong Jiang[a,b] and Keming Yu[b,c]

[a]Department of Statistics, Donghua University, Shanghai, China; [b]College of Science, Anqing Normal University, Anqing, China;
[c]Brunel University London, Uxbridge, UK

ABSTRACT
Single-index quantile regression (QR) models can avoid the curse of dimensionality in nonparametric problems by assuming that the response is only related to a single linear combination of the covariates. Like the standard parametric or nonparametric QR whose estimated curves may cross, the single-index QR can also suffer quantile crossing, leading to an invalid distribution for the response. This issue has attracted considerable attention in the literature in the recent year. In this article, we consider single-index models, develop methods for QR that guarantee noncrossing quantile curves, and extend the methods and results to composite quantile regression. The asymptotic properties of the proposed estimators are derived and their advantages over existing methods are explained. Simulation studies and a real data application are conducted to illustrate the finite sample performance of the proposed methods.

ARTICLE HISTORY
Received June 2021
Accepted November 2021

KEYWORDS
Composite quantile regression; Crossing quantile curves; Quantile regression; Single-index model

1. Introduction

Quantile regression (QR) (Koenker 2005) has been receiving increasing attention in econometrics and statistics. Its advantages over the other estimation methods have been well investigated. As an alternative to the least-square regression, QR provides a complete statistical analysis of the stochastic relationships among variables. Specifically, QR is used to estimate the τth conditional quantile of the response variable. Let

$$Q_{\mathbf{Y}}(\tau|\mathbf{x}) = Q_{\mathbf{Y}}(\tau|\mathbf{X}=\mathbf{x}) = \inf\{y : P(\mathbf{Y} \le y|\mathbf{X}=\mathbf{x}) \ge \tau\},$$

denote the τth conditional quantile of the response $\mathbf{Y}$ given a p-dimensional vector of covariates $\mathbf{X}=\mathbf{x}$. In short, $Q_{\mathbf{Y}}(\tau|\mathbf{x})$ is the smallest real value such that the probability of obtaining smaller values of $\mathbf{Y}$ is at least τ. The traditional QR is concerned with the estimation of the τth conditional QR of $\mathbf{Y}$ for a given $\mathbf{x}$, which often sets a linear model as follows:

$$Q_{\mathbf{Y}}(\tau|\mathbf{x}) = \mathbf{x}^\top \gamma_{01,\tau},$$

where $\gamma_{01,\tau}$ is a p-dimensional vector of unknown parameters. Because this model has a strict linearity assumption and lacks the flexibility to deal with various nonlinearities present in some datasets, several authors have considered the completely flexible nonparametric estimation of the conditional quantiles; see Chaudhuri (1991), Yu and Jones (1998), Takeuchi et al. (2006), Li and Racine (2008), Guerre and Sabbah (2012), Kong et al. (2013), Huang and Nguyen (2018), and among others.

As in the case of conditional mean regression, the nonparametric estimation of $Q_{\mathbf{Y}}(\tau|\mathbf{x})$ suffers from the so-called curse of dimensionality. One way to solve this is to consider the single-index model (Ichimura 1993). Specifically, given $\tau \in (0,1)$, we model the τth conditional quantile of the response variable $\mathbf{Y}$ at the multivariate covariate vector of values $\mathbf{X}=\mathbf{x}$ with a single-index structure,

$$Q_{\mathbf{Y}}(\tau|\mathbf{x}) = g_{0,\tau}(\mathbf{x}^\top \gamma_{01,\tau}), \tag{1.1}$$

where $g_{0,\tau}(\cdot)$ is the unknown link function and $\gamma_{01,\tau}$ is the unknown parameter vector. For identifiability, one imposes certain conditions on $\gamma_{01,\tau}$, and it is often assumed that $\|\gamma_{01,\tau}\|_2 = 1$ and the first nonzero element of $\gamma_{01,\tau}$ is positive. Note that $\|\gamma_{01,\tau}\|_2 = 1$; it means that the true value of $\gamma_{01,\tau}$ is the boundary point on the unit sphere, and hence $g_{0,\tau}(\mathbf{x}^\top\gamma_{01,\tau})$ does not have a derivative at the point $\gamma_{01,\tau}$ (Zhu and Xue 2006). To overcome this difficulty, we make the following adjustment by using the "delete-one-component" method proposed by Yu and Ruppert (2002). We assume that $\gamma_{01,\tau} = (1, \gamma_{0,\tau}^\top)^\top$ with $\gamma_{0,\tau} \in R^{p-1}$. Christou and Akritas (2016) also used this adjustment. This means that we transform a restricted problem of estimating $\gamma_{01,\tau}$ to an unrestricted problem.

The single-index model has the following advantages: (i) the single-index in the link function projects multivariate covariates onto a one-dimensional variate, which effectively overcomes the "curse of dimensionality"; (ii) the unspecified link function allows model flexibility and thus has a lower risk of misspecification; and (iii) the interpretation of covariate effects is easy because of the linear structure of the index. Therefore, single-index quantile regression (1.1) has received extensive attention in the literature in the recent years; see Chaudhuri et al. (1997), Wu et al. (2010), Oh et al. (2011), Kong and Xia (2012), Lv et al. (2015), Christou and Akritas (2016), Jiang et al. (2016),

CONTACT Keming Yu keming.yu@brunel.ac.uk Department of Mathematics, Brunel University London, Uxbridge, Middlesex, UB8 3PH, UK.
Both authors have equal contribution.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/UBES.

Ma and He (2016), Tang et al. (2018), Jiang and Yu (2020), and among others.

Population conditional quantile functions cannot cross each other for different quantile orders; however, the estimated regression curves often violate this, which can be very challenging for interpretation and further analysis. Considerable attention in the literature has been devoted to studying the noncrossing estimation of a variety of quantile regression models for years. For linear and nonlinear quantile estimation, He (1997) restricted the possible solution space of the response distribution to location-scale changes of a base distribution to obtain noncrossing curves. Wu and Liu (2009) had proposed a stepwise procedure to ensure noncrossing estimated curves. Santos and Kneib (2020) considered a flexible Bayesian quantile regression model with Gaussian process adjustment to achieve a noncrossing property. For nonparametric noncrossing quantile estimation, Yu and Jones (1998) had suggested a double kernel smoothing method with a minor modification of this estimate in a second step so that the corresponding quantile curves are monotonic. Hall et al. (1999) had proposed an adjusted Nadaraya-Watson estimator, which modifies the weights so that the resulting estimate of the conditional distribution function is monotonic. Dette and Volgushev (2008) considered noncrossing estimates of quantile curves using a simultaneous inversion and isotonization of an estimate of the conditional distribution function. Bondell et al. (2010) developed a method for noncrossing quantile regression (NCQR) curve estimation using spline based constraints. Chernozhukov et al. (2010) proposed estimating noncrossing quantile curves via a monotonic rearrangement of the original nonmonotonic function. Schnabel and Eilers (2013) developed a quantile sheet that contained the quantile level as an argument in the quantile curve. Andriyana et al. (2018) extended the methods of He (1997), Wu and Liu (2009), Bondell et al. (2010), and Schnabel and Eilers (2013) to prevent the crossing of the estimated quantile curves in a setting of varying-coefficient modeling. Kim et al. (2019) derived a monotonic constraint tree model for quantile regression. Chen et al. (2021) had studied the noncrossing problem for semiparametric multi-index quantile regression while requiring that the unknown parameter not depend on quantile level τ; however, this cannot guarantee quantile noncrossing under model (1.1). This article aims to develop a new estimation method for model (1.1) to avoid quantile crossing.

Most existing methods for addressing quantile crossing are based on a restricted version of the regression quantile or set a monotonic constraint. They cannot be directly applied to single-index quantile regression, because there are both parametric ($\gamma_{01,\tau}$) and nonparametric ($g_{0,\tau}(\cdot)$) parts in model (1.1). The innovation of this article is the simple kernel estimation of the nonparametric part ($g_{0,\tau}(\cdot)$) in model (1.1) to avoid quantile crossing. Therefore, the restrictive estimator of an unknown parameter ($\gamma_{01,\tau}$) in model (1.1) is valid, where the restrictive condition guarantees the noncrossing estimator of the quantile function $Q_{\mathbf{Y}}(\tau|\mathbf{x})$. The method does not impose any assumptions on $g_{0,\tau}(\cdot)$ except for some continuity and smoothness conditions. Furthermore, we study its extension called composite quantile regression (CQR) proposed by Zou and Yuan (2008). Jiang et al. (2016) extended the CQR method to a single-index model. Christou and Akritas (2016) also developed a kernel estimation of the nonparametric part for model (1.1). However, their method needs to estimate $Q_{\mathbf{Y}}(\tau|\mathbf{x})$ first, which may face the "curse of dimensionality". More importantly, the estimator of the nonparametric part cannot avoid quantile crossing. Therefore, the restrictive estimator of unknown parameters in model (1.1) by their method is not necessarily effective.

Overall, this study offers a novel approach and makes the following key contributions:

i. We develop methods to prevent the crossing of the estimated quantile curves in single-index modeling. The proposed estimation methods are easy to implement for computation.
ii. We study the CQR method which can improve the efficiency.

The article is organized as follows. In Section 2, we introduce the NCQR and derive the asymptotic theories of new estimations for model (1.1). In Section 3, a noncrossing single-index CQR is proposed. Simulation examples and a real data application are given in Section 4 to illustrate the proposed procedures. Final remarks are given in Section 5. All technical proofs are deferred to the supplementary material.

2. Noncrossing Single-Index Quantile Regression

2.1. The Noncrossing Estimator of $g_{0,\tau}(\cdot)$

Let $\{Y_i, \mathbf{X}_i\}_{i=1}^n$ with $\mathbf{X}_i \in \mathbf{D} \subset R^p$ be an independent identically distributed (iid) sample from $(Y, \mathbf{X})$. The classical kernel estimation is based on minimizing the following local linear sample with respect to (a, b) to obtain the estimator $\hat{g}_\tau(u|\gamma_{0,\tau}) = \hat{a}$ of $g_{0,\tau}(u|\gamma_{0,\tau})$, where

$$(\hat{a}, \hat{b}) = \arg\min_{(a,b)} \sum_{i=1}^n \rho_\tau\left(Y_i - a - b(\mathbf{X}_i^\top \gamma_{01,\tau} - u)\right) K\left(\frac{\mathbf{X}_i^\top \gamma_{01,\tau} - u}{h}\right),$$

where $\rho_\tau(r) = \tau r - rI(r < 0)$ is the check loss function, $I(\cdot)$ is the indicator function, $K(\cdot)$ is the kernel weight function and h is called bandwidth, which is used to control the amount of smoothing. See the details from Wu et al. (2010) and among others. However, from the numerical examples of Section 4, this estimation method may not avoid quantile crossing.

Let $F(\mathbf{Y}|\mathbf{X}^\top\gamma_{01,\tau} = \mathbf{U}_{0,\tau})$ be the conditional distribution function of $\mathbf{Y}$ given $\mathbf{X}^\top\gamma_{01,\tau} = \mathbf{U}_{0,\tau}$; then, $g_{0,\tau}(\mathbf{U}_{0,\tau}|\gamma_{0,\tau})$ in model (1.1) can be regarded to satisfy

$$F(\cdot|\mathbf{X}^\top\gamma_{01,\tau} = \mathbf{U}_{0,\tau}) = \tau.$$

Then, instead of the "check-function"-based classical kernel quantile regression estimation above, we estimate $g_{0,\tau}(\mathbf{U}_{0,\tau}|\gamma_{0,\tau})$ via the inverse function of the conditional distribution $F(\mathbf{Y}|\mathbf{X}^\top \gamma_{01,\tau} = \mathbf{U}_{0,\tau})$. We will show that by properly selecting the kernel function, this alternative single-index quantile regression estimation can automatically produce an estimator of $g_{0,\tau}(\mathbf{U}_{0,\tau}|\gamma_{0,\tau})$ over the design set. This estimator is a monotonic function of τ and avoids quantile crossing.

In fact, $F(y|\mathbf{X}^\top\gamma_{01,\tau} = \mathbf{U}_{0,\tau})$ is unknown, and its well-known kernel estimator is

$$\hat{F}(y|\mathbf{X}^\top\gamma_{01,\tau} = u) = \frac{\sum_{j=1}^n K_h(u - U_{j,\tau}) I(Y_j \le y)}{\sum_{j=1}^n K_h(u - U_{j,\tau})},$$

where $U_{j,\tau} = \mathbf{X}_j^\top \gamma_{01,\tau}$ and $K_h(\cdot) = K(\cdot/h)/h$. We replace the indicator function $I(\cdot)$ by a smoothing distribution function as

$$\hat{F}(y|\mathbf{X}^\top \gamma_{01,\tau} = u) = \frac{\sum_{j=1}^n K_h(u-U_{j,\tau})\Omega\left(\frac{y-Y_j}{h_1}\right)}{\sum_{j=1}^n K_h(u-U_{j,\tau})}, \qquad (2.1)$$

with

$$\Omega(y) = \int_{-\infty}^{y} W(v)dv,$$

and the associated bandwidth h_1, where $W(\cdot)$ could be any symmetric probability density function.

It is well known that the selection of both bandwidths h and h_1 are important in nonparametric kernel smoothing methods, but the selection of kernel functions $K(\cdot)$ and $W(\cdot)$ are not as important as h and h_1. Specifically, we choose $W(\cdot)$ as the uniform kernel: $W(v) = 1/2I(|v| \le 1)$, for simple calculation. This affords the quantile function $g_\tau \equiv g_\tau(u|\gamma_{0,\tau})$ to satisfy

$$\begin{aligned}
\frac{1}{h_1}\int_{-\infty}^{g_\tau} W\left(\frac{Y_j - v}{h_1}\right)dv
&= \frac{1}{2h_1}\int_{-\infty}^{g_\tau} I(Y_j - h_1 \le v \le Y_j + h_1)dv \\
&= \frac{1}{2h_1} I(g_\tau \ge Y_j - h_1)\int_{Y_j-h_1}^{\min(g_\tau, Y_j+h_1)} dv \\
&= \frac{1}{2h_1}\big\{(g_\tau - Y_j + h_1) - (g_\tau - Y_j - h_1)I(g_\tau \ge Y_j + h_1) \\
&\quad - (g_\tau - Y_j + h_1)I(g_\tau \le Y_j - h_1)\big\},
\end{aligned}$$

and that $\hat{g}_\tau(u|\gamma_{0,\tau})$, for each u, is the value of $g_\tau(u|\gamma_{0,\tau})$, which is the solution of

$$\begin{aligned}
g_\tau(u|\gamma_{0,\tau}) =& (2\tau - 1)h_1 + \frac{\sum_{j=1}^n Y_j K_h(u-U_{j,\tau})}{\sum_{j=1}^n K_h(u-U_{j,\tau})} \\
&+ \frac{\sum_{j=1}^n \{g_\tau(u|\gamma_{0,\tau}) - Y_j - h_1\}K_h(u-U_{j,\tau})I\{Y_j \le g_\tau(u|\gamma_{0,\tau}) - h_1\}}{\sum_{j=1}^n K_h(u-U_{j,\tau})} \\
&+ \frac{\sum_{j=1}^n \{g_\tau(u|\gamma_{0,\tau}) - Y_j + h_1\}K_h(u-U_{j,\tau})I\{Y_j \ge g_\tau(u|\gamma_{0,\tau}) + h_1\}}{\sum_{j=1}^n K_h(u-U_{j,\tau})}.
\end{aligned} \qquad (2.2)$$

When $\gamma_{01,\tau}$ is independent of τ, we can rewrite $U_{j,\tau} = \mathbf{X}_j^\top \gamma_{01,\tau}$ as U_j. Then, taking the first derivative over τ on both sides of Equation (2.2), we can prove the following lemma:

$$\frac{\partial g_\tau(u|\gamma_{0,\tau})}{\partial \tau} = \frac{2h_1 \sum_{j=1}^n K_h(u - U_j)}{\sum_{j=1}^n K_h(u - U_j)I\{g_\tau(u|\gamma_{0,\tau}) - h_1 < Y_j < g_\tau(u|\gamma_{0,\tau}) + h_1\}} > 0,$$

where $\partial g_\tau(u|\gamma_{0,\tau})/\partial\tau$ is the first derivative of $g_\tau(u|\gamma_{0,\tau})$ with respect to τ. That is, the estimated $g_{0,\tau}(u|\gamma_{0,\tau})$ by this new method is a monotonic function of τ for all u.

By (2.2) and a $\sqrt{nh}$-consistent initial estimator $\tilde{g}_\tau(u)$ of $g_{0,\tau}(u|\gamma_{0,\tau})$, which can be obtained by Wu et al. (2010), we can propose the estimator of $g_{0,\tau}(u|\gamma_{0,\tau})$ by

$$\begin{aligned}
\hat{g}_\tau(u|\gamma_{0,\tau}) =& (2\tau - 1)h_1 + \frac{\sum_{j=1}^n Y_j K_h(u-U_{j,\tau})}{\sum_{j=1}^n K_h(u-U_{j,\tau})} \\
&+ \frac{\sum_{j=1}^n \{\tilde{g}_\tau(u) - Y_j - h_1\}K_h(u-U_{j,\tau})I\{Y_j \le \tilde{g}_\tau(u) - h_1\}}{\sum_{j=1}^n K_h(u-U_{j,\tau})} \\
&+ \frac{\sum_{j=1}^n \{\tilde{g}_\tau(u) - Y_j + h_1\}K_h(u-U_{j,\tau})I\{Y_j \ge \tilde{g}_\tau(u) + h_1\}}{\sum_{j=1}^n K_h(u-U_{j,\tau})}.
\end{aligned} \qquad (2.3)$$

2.2. The Noncrossing Estimator of $Q_{\mathbf{Y}}(\tau|\mathbf{x})$

The main purpose of this article is to study $Q_{\mathbf{Y}}(\tau|\mathbf{x}) = g_{0,\tau}(\mathbf{x}^\top \gamma_{01,\tau})$ without the crossing problem. Suppose we want to estimate quantile functions $Q_{\mathbf{Y}}(\tau_t|\mathbf{x})$ simultaneously at $0 < \tau_1 < \cdots < \tau_q < 1$.

When $\gamma_{01,\tau}$ is independent of τ, we first obtain the estimator $\hat{\gamma}_1$ of $\gamma_{01,\tau}$ by the method in Wu et al. (2010), Kong and Xia (2012), or Christou and Akritas (2016) ($\tau = 0.5$ is recommended). Then, as discussed in Section 2.1, $\hat{Q}_{\mathbf{Y}}(\tau_t|\mathbf{x}) = \hat{g}_{\tau_t}(\mathbf{x}^\top \hat{\gamma}_1)$ is a noncrossing estimator of $Q_{\mathbf{Y}}(\tau_t|\mathbf{x})$ for $t = 1, \ldots, q$, where $\hat{g}_{\tau_t}(\cdot)$ is defined in (2.3) with $u = \mathbf{x}^\top \hat{\gamma}_1$ and $U_{j,\tau} = \mathbf{X}_j^\top \hat{\gamma}_1$.

For the single-index quantile regression model (1.1), $\gamma_{01,\tau}$ should depend on τ. In this case, we need some restrictions on the estimator of $\gamma_{01,\tau}$ to obtain noncrossing curves. We estimate $\gamma_0(\tau) = (\gamma_{0,\tau_1}^\top, \ldots, \gamma_{0,\tau_q}^\top)^\top$ by the following noncrossing restriction:

$$\begin{aligned}
\hat{\gamma}(\tau) = \;& \arg\min_{\gamma_{\tau_t}} \frac{1}{n}\sum_{t=1}^q \sum_{i=1}^n \rho_{\tau_t}\left(Y_i - \hat{g}_{\tau_t}(\mathbf{X}_i^\top \gamma_{1,\tau_t}|\gamma_{\tau_t})\right) \\
s.t. \;& \hat{g}_{\tau_t}(\mathbf{x}^\top \gamma_{1,\tau_t}|\gamma_{\tau_t}) \ge \hat{g}_{\tau_{t-1}}(\mathbf{x}^\top \gamma_{1,\tau_{t-1}}|\gamma_{\tau_{t-1}}),\ \forall \mathbf{x} \in \mathbf{D},
\end{aligned} \qquad (2.4)$$

where $\hat{\gamma}(\tau) = (\hat{\gamma}_{\tau_1}^\top, \ldots, \hat{\gamma}_{\tau_q}^\top)^\top$ and $\gamma_{1,\tau_t} = (1, \gamma_{\tau_t}^\top)^\top$ with $\gamma_{\tau_t} \in R^{p-1}$, for $t = 1, \ldots, q$. According to the estimator $\hat{\gamma}(\tau)$ of $\gamma_0(\tau)$ by (2.4), we can have $\hat{Q}_{\mathbf{Y}}(\tau_t|\mathbf{x}) = \hat{g}_{\tau_t}(\mathbf{x}^\top \hat{\gamma}_{1,\tau_t})$, $t = 1, \ldots, q$, which are noncrossing by the constraint in (2.4). Moreover, because of the noncrossing of $\{\hat{g}_{\tau_t}(\cdot)\}_{t=1}^q$, at least one solution satisfying the restriction condition in (2.4) is $\hat{\gamma}_{\tau_1} = \cdots = \hat{\gamma}_{\tau_q}$, which are independent of τ.

2.3. Asymptotic Normality of Estimators

Let $F(\cdot)$ and $f(\cdot)$ be the cumulative distribution function and density function of $\mathbf{Y}$, respectively. Denote by $f_{\mathbf{U}_{0,\tau}}(\cdot)$ the marginal density function of $\mathbf{U}_{0,\tau}$, $f'_{\mathbf{U}_{0,\tau}}(\cdot)$ is the derivative of $f_{\mathbf{U}_{0,\tau}}(\cdot)$, $f(\cdot|u)$ is the conditional density of $\mathbf{Y}$ given u, $\mu_2 = \int u^2 K(u)du$ and $\nu_0 = \int K^2(u)du$. Moreover, we make the following assumptions.

Condition 1: The density function of $\mathbf{X}^\top \gamma_1$ is positive and uniformly continuous for γ_1 in a neighborhood of $\gamma_{01,\tau}$. Furthermore, $f_{\mathbf{U}_{0,\tau}}(u_{0,\tau})$ has a continuous and bounded second derivative with respect to $u_{0,\tau} = \mathbf{x}^\top \gamma_{01,\tau}$. For a fixed value of $u_{0,\tau}, f_{\mathbf{U}_{0,\tau}}(u_{0,\tau}) > 0$.

Condition 2: The function $g_{0,\tau}(\cdot)$ has a continuous and bounded second derivative.

Condition 3: The conditional cumulative distribution function of $\mathbf{Y}$ given u, $F(y|u)$, has a uniformly continuous second-order partial derivative function with respect to $\mathbf{Y}$ and u. The conditional density of $\mathbf{Y}$ given u, $f(y|u)$, is continuous in u for each y, and $0 < f(g_{0,\tau}(u)|u) \le c$, where c is a positive constant. Moreover, there exist positive constants ϵ and δ and a positive function $G(y|u)$ such that

$$\sup_{|u_n - u| \le \epsilon} f(y|u_n) \le G(y|u),\quad \int |\rho'_\tau(y - g_{0,\tau}(u))|^{2+\delta} G(y|u)dy < \infty,$$
$$\int \{\rho_\tau(y - t) - \rho_\tau(y) - \rho'_\tau(y)t\}^2 G(y|u)dy = o(t^2) \ \ as \ \ t \to 0.$$

Condition 4: Define $\varphi_\tau(t|\mathbf{x}) = E\big\{\rho_\tau\left(Y_i - g_{0,\tau}(\mathbf{x}^\top \gamma_{01,\tau}) + t\right) |\mathbf{X} = \mathbf{x}\big\}$, for which the expectation and differentiation can be interchanged, and let the first and second derivatives of $\varphi_\tau(t|\mathbf{x})$ with respect to t exist.

Condition 5: The kernel $K(\cdot)$ is a symmetric, bounded, and compactly supported density function.

Condition 6: The following expectations exist,

$$\mathbf{S}_0 = E\left[\{g'_{0,\tau}(\mathbf{X}^\top\gamma_{01,\tau})\}^2\tilde{\mathbf{X}}_{-1}\tilde{\mathbf{X}}_{-1}^\top\right],$$
$$\mathbf{S}_1 = E\left[f(g_{0,\tau}(\mathbf{X}^\top\gamma_{01,\tau})|\mathbf{X}^\top\gamma_{01,\tau})\{g'_{0,\tau}(\mathbf{X}^\top\gamma_{01,\tau})\}^2\tilde{\mathbf{X}}_{-1}\tilde{\mathbf{X}}_{-1}^\top\right],$$

where $\tilde{\mathbf{X}}_{-1} = \mathbf{X}_{-1} - E(\mathbf{X}_{-1}|\mathbf{X}^\top\gamma_{01,\tau})$, and $\mathbf{X}_{-1}$ is the $(p-1)$-dimensional vector consisting of coordinates $2,\dots,p$ of $\mathbf{X}$. Moreover, $\mathbf{S}_1$ is assumed to be nonsingular.

Remark 2.1. Conditions 1–6 are standard conditions, which are commonly used in single-index regression; see Wu et al. (2010) and Christou and Akritas (2016). Condition 1 guarantees the existence of any ratio terms with the density appearing as part of the denominator. Condition 2 is a common assumption for the link function. Condition 3 is weaker than the Lipschitz continuity of the function $\rho'_\tau(\cdot)$. Condition 4 imposes smoothness conditions on $\varphi_\tau(\cdot|\mathbf{x})$, since $\rho_\tau(\cdot)$ is actually not differentiable at 0. Condition 5 simply requires that the kernel function is a proper density with a finite second moment that is required for the asymptotic variance of estimators. Condition 6 ensures the existence of the asymptotic variance of estimators.

Theorem 2.1. Under Conditions 1–6, $nh^9 \to 0$, $nh^5h_1^4 \to 0$, $nh \to \infty$, $n \to \infty$, $h_1 \to 0$, $\sqrt{nh}\min_t(\tau_t - \tau_{t-1}) \to \infty$ with $0 < \tau_1 < \cdots < \tau_q < 1$, and the initial estimator $\tilde{g}_\tau(\cdot)$ in (2.3) is $\sqrt{nh}$-consistent. Then for any $\mathbf{z} \in R^{pq}$, we have

$$\left|P\left[\sqrt{n}\left\{\hat{\gamma}(\tau) - \gamma_0(\tau)\right\} \le \mathbf{z}\right] - P\left[\sqrt{n}\{\bar{\gamma}(\tau) - \gamma_0(\tau)\} \le \mathbf{z}\right]\right| \to 0,$$

where $\hat{\gamma}(\tau)$ is defined in (2.4) and $\bar{\gamma}(\tau) = (\bar{\gamma}_{\tau_1}^\top,\dots,\bar{\gamma}_{\tau_q}^\top)^\top$ is the unconstrained estimator with

$$\bar{\gamma}_{\tau_t} = \arg\min_{\gamma_{\tau_t}} \frac{1}{n}\sum_{i=1}^n \rho_{\tau_t}\left(Y_i - \hat{g}_{\tau_t}(\mathbf{X}_i^\top\gamma_{1,\tau_t}|\gamma_{\tau_t})\right), \quad t = 1,\dots,q. \tag{2.5}$$

Moreover, for any $\tau \in (\tau_1,\dots,\tau_q)$, we have

$$\sqrt{n}(\hat{\gamma}_\tau - \gamma_{0,\tau}) \xrightarrow{L} \mathcal{N}\left(0, \tau(1-\tau)\mathbf{S}_1^{-1}\mathbf{S}_0\mathbf{S}_1^{-1}\right),$$

where $\xrightarrow{L}$ stands for convergence in the distribution.

Theorem 2.2. Under the same conditions as in Theorem 2.1, then for an interior point u,

$$\sqrt{nh}\left\{\hat{g}_\tau(u|\hat{\gamma}_\tau) - g_{0,\tau}(u) - \alpha_\tau(u)\right\} \xrightarrow{L} \mathcal{N}(0, \Sigma_\tau(u)),$$

where

$$\alpha_\tau(u) = f^{-1}(g_{0,\tau}(u)|u)\left\{\frac{1}{2}h^2\mu_2\beta_\tau(u) + \frac{1}{6}h_1^2\frac{\partial^2 F(y|u)}{\partial y^2}\Big|_{y=g_{0,\tau}(u)}\right\},$$
$$\beta_\tau(u) = \frac{\partial^2 F(g_{0,\tau}(u)|u)}{\partial u^2} + 2f_{\mathbf{U}_{0,\tau}}^{-1}(u)f'_{\mathbf{U}_{0,\tau}}(u)\frac{\partial F(g_{0,\tau}(u)|u)}{\partial u},$$
$$\Sigma_\tau(u) = \frac{\nu_0\{\tau(1-\tau) - 1/3h_1 f(g_{0,\tau}(u)|u)\}}{f_{\mathbf{U}_{0,\tau}}(u)f^2(g_{0,\tau}(u)|u)}.$$

Under Conditions 1 and 3 where $f(g_{0,\tau}(u)|u) > 0$ and $f_{\mathbf{U}_{0,\tau}}(u) > 0$, the asymptotic variance $\Sigma_\tau(u)$ is smaller than $\nu_0\tau(1-\tau)/\{f_{\mathbf{U}_{0,\tau}}(u)f^2(g_{0,\tau}(u)|u)\}$, the asymptotic variance of existing estimates; see Wu et al. (2010) and Christou and Akritas (2016). This asymptotically smaller variance of the proposed estimators than that of existing estimators may partially explain why it can avoid quantile crossing. Moreover, the bias of the proposed method is different from that of Wu et al. (2010) or Christou and Akritas (2016). However, it is difficult to compare their differences.

With the results of Theorems 2.1 and 2.2, it is straightforward to obtain a conditional quantile estimator $\hat{Q}_{\mathbf{Y}}(\tau|\mathbf{x}) = \hat{g}(\mathbf{x}^\top\hat{\gamma}_{1,\tau}|\hat{\gamma}_\tau)$ of $Q_{\mathbf{Y}}(\tau|\mathbf{x})$ in (1.1), where $\hat{\gamma}_{1,\tau} = (1,\hat{\gamma}_\tau^\top)^\top$. The asymptotic normality of $\hat{Q}_{\mathbf{Y}}(\tau|\mathbf{x})$ is as follows.

Theorem 2.3. Under the same conditions as in Theorem 2.2 and $nh^4 \to \infty$, then for an interior point $\mathbf{x}$,

$$\sqrt{nh}\left\{\hat{Q}_{\mathbf{Y}}(\tau|\mathbf{x}) - Q_{\mathbf{Y}}(\tau|\mathbf{x}) - \alpha_\tau(\mathbf{x}^\top\gamma_{01,\tau})\right\} \xrightarrow{L} \mathcal{N}\left(0, \Sigma_\tau(\mathbf{x}^\top\gamma_{01,\tau})\right).$$

2.4. Selection of Bandwidths

In this section, we focus on how to choose the bandwidths h and h_1. Theorem 2.2 implies that the mean squared error (MSE) of $\hat{g}_\tau(u|\hat{\gamma}_\tau)$ is

$$\begin{aligned}\mathrm{MSE}(\hat{g}_\tau(u|\hat{\gamma}_\tau)) &= \alpha_\tau^2(u) + (nh)^{-1}\Sigma_\tau(u)\\ &= C_1h^4 + C_2h^2h_1^2 + C_3h_1^4 + C_4(nh)^{-1}\\ &\quad -C_5h_1(nh)^{-1},\end{aligned} \tag{2.6}$$

where $C_1 = 1/4f^{-2}(g_{0,\tau}(u)|u)\mu_2^2\beta_\tau^2(u)$, $C_2 = 1/6f^{-2}(g_{0,\tau}(u)|u)\mu_2\beta_\tau(u)\partial^2F(y|u)/\partial y^2|_{y=g_{0,\tau}(u)}$, $C_3 = 1/36f^{-2}(g_{0,\tau}(u)|u)\partial^2F(y|u)/\partial y^2|_{y=g_{0,\tau}(u)}$, $C_4 = \nu_0\tau(1-\tau)f_{\mathbf{U}_{0,\tau}}^{-1}(u)f^{-2}(g_{0,\tau}(u)|u)$ and $C_5 = 1/3\nu_0 f_{\mathbf{U}_{0,\tau}}^{-1}(u)f^{-1}(g_{0,\tau}(u)|u)$ are all constants.

The first-order conditions required for minimization of (2.6) are

$$\frac{\partial\mathrm{MSE}(\hat{g}_\tau(u|\hat{\gamma}_\tau))}{\partial h} = 4C_1h^3 + 2C_2hh_1^2 - C_4(nh^2)^{-1} + C_5h_1(nh^2)^{-1} = 0, \tag{2.7}$$

$$\frac{\partial\mathrm{MSE}(\hat{g}_\tau(u|\hat{\gamma}_\tau))}{\partial h_1} = 2C_2h^2h_1 + 4C_3h_1^3 - C_5(nh)^{-1} = 0. \tag{2.8}$$

From (2.7) and (2.8), one can see that h_1 must have an order smaller than that of h. Furthermore, assume $h = cn^{-\theta}$ and $h_1 = c_1n^{-\theta_1}$. Then we must have $\theta_1 > \theta$. One can try to use $\theta_1 = \theta$ or $\theta_1 < \theta$, which will lead to a contradiction. Therefore, from (2.7) we have $\theta = 1/5$, and substituting it into (2.8) yields $\theta_1 = 2/5$. Thus, the optimal bandwidths are $h = cn^{-1/5}$ and $h_1 = c_1n^{-2/5}$.

The above analysis provides only the optimal rate. Although in principle one can compute plug-in bandwidths based on (2.6), this will not be feasible in applied settings. In practice, the bandwidths h and h_1 can be selected by minimizing the following cross-validation function (Li et al. 2013):

$$\mathrm{CV}(h,h_1) = \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left\{I(Y_i \le Y_j) - \hat{F}_{-i}(Y_j|\hat{U}_{i,\tau})\right\}^2, \tag{2.9}$$

where $\hat{F}_{-i}(\cdot)$ is the "leave-one-out" kernel estimator by (2.1) and $\hat{U}_{i,\tau} = \mathbf{X}_i^\top\hat{\gamma}_{1,\tau}$. Moreover, (2.9) can be easily solved by using the "npcdistbw" function in the R package "np."

2.5. The Algorithm

In this section, we introduce the algorithm of the methods proposed in Sections 2.1 and 2.2.

To ensure that the estimated quantile curves do not cross, we adopt the stepwise procedure (Wu and Liu 2009). It starts by estimating a particular quantile function; $\tau = 0.5$ was advised by Wu and Liu (2009). In the next step, we estimate the next (higher or lower order) quantile. In the upward step, we add a constraint so that the estimated higher order quantile curve exceeds the preceding estimated quantile curve. In the downward step, we put a constraint such that the estimated lower order quantile curve does not exceed the preceding estimated quantile curve. The procedure continues by moving to the next quantile order.

Moreover, to obtain the estimator $\hat{\gamma}(\tau)$ by minimizing (2.4), we use a local linear approximation of $\hat{g}_{\tau_t}(\mathbf{X}_i^\top \gamma_{1,\tau_t}|\gamma_{\tau_t})$ around a $\sqrt{n}$-consistent initial value $\tilde{\gamma}_{\tau_t}$ of γ_{0,τ_t}. This yields

$$\hat{g}_{\tau_t}(\mathbf{X}_i^\top \gamma_{1,\tau_t}|\gamma_{\tau_t}) \approx \hat{g}_\tau(\mathbf{X}_i^\top \tilde{\gamma}_{1,\tau_t}|\tilde{\gamma}_{\tau_t}) + \hat{g}'_\tau(\mathbf{X}_i^\top \tilde{\gamma}_{1,\tau_t}|\tilde{\gamma}_{\tau_t})\mathbf{X}_{i,-1}^\top(\gamma_{\tau_t} - \tilde{\gamma}_{\tau_t}),$$

where $\tilde{\gamma}_{1,\tau_t} = (1, \tilde{\gamma}_{\tau_t}^\top)^\top$ and $\hat{g}'_\tau(u) = \partial \hat{g}_\tau(u)/\partial u$.

The steps of the proposed procedure are summarized as follows.

Step 0: Obtain initial estimators of $\gamma_{0,\tau}$ and $g_{0,\tau}(\cdot)$.

The initial estimators $\tilde{\gamma}_\tau$ of $\gamma_{0,\tau}$ and $\tilde{g}_\tau(\mathbf{X}_i^\top \tilde{\gamma}_{1,\tau})$ of $g_{0,\tau}(\mathbf{X}_i^\top \tilde{\gamma}_{1,\tau})$ for $\tau \in (\tau_1, \ldots, \tau_q)$, can be obtained by the method in Wu et al. (2010), Kong and Xia (2012), or Christou and Akritas (2016), which all satisfy the condition of $\sqrt{n}$ and $\sqrt{nh}$ consistency for $\gamma_{0,\tau}$ and $g_{0,\tau}(\cdot)$, respectively. In our numerical study, we obtain all initial estimators by methods in Wu et al. (2010). Note that $\gamma_{01,\tau} = (1, \gamma_{0,\tau}^\top)^\top$, we modify the estimator of $\gamma_{01,\tau}$ in Wu et al. (2010) by setting its first component equal to 1.

Step 1: Estimate $\gamma_{0,\tau}$ with $\tau = 0.5$ (unconstrained estimation).

The median quantile estimator is given by minimizing the following objective function with respect to γ:

$$\begin{aligned}&\frac{1}{n}\sum_{i=1}^n \rho_{0.5}\left(Y_i - \hat{g}_{0.5}(\mathbf{X}_i^\top \tilde{\gamma}_{1,0.5}|\tilde{\gamma}_{0.5}) - \hat{g}'_{0.5}(\mathbf{X}_i^\top \tilde{\gamma}_{1,0.5}|\tilde{\gamma}_{0.5})\mathbf{X}_{i,-1}^\top(\gamma - \tilde{\gamma}_{0.5})\right)\\ &= \frac{1}{n}\sum_{i=1}^n \rho_{0.5}\left(Y_{i,0.5}^* - \hat{g}'_{0.5}(\mathbf{X}_i^\top \tilde{\gamma}_{1,0.5}|\tilde{\gamma}_{0.5})\mathbf{X}_{i,-1}^\top \gamma\right),\end{aligned} \tag{2.10}$$

where $Y_{i,\tau}^* = Y_i - \hat{g}_\tau(\mathbf{X}_i^\top \tilde{\gamma}_{1,\tau}|\tilde{\gamma}_\tau) + \hat{g}'_\tau(\mathbf{X}_i^\top \tilde{\gamma}_{1,\tau}|\tilde{\gamma}_\tau)\mathbf{X}_{i,-1}^\top \tilde{\gamma}_\tau$. Note that (2.10) can be easily solved by using the "rq" function in the R package "quantreg".

Step 2: Complete up ($\gamma_{0,\tau}$).

Starting from $\tau_t = 0.5$, the next higher order $\tau_{t+1} > \tau_t$ is obtained from the following constrained minimization problem:

$$\begin{aligned}\hat{\gamma}_{\tau_{t+1}} &= \arg\min_{\gamma_{\tau_{t+1}}} \frac{1}{n}\sum_{i=1}^n \rho_{\tau_t}\left(Y_{i,\tau_{t+1}}^* - \hat{g}'_{\tau_{t+1}}(\mathbf{X}_i^\top \tilde{\gamma}_{1,\tau_{t+1}}|\tilde{\gamma}_{\tau_{t+1}})\mathbf{X}_{i,-1}^\top \gamma_{\tau_{t+1}}\right)\\ \text{s.t.} \quad & \hat{g}_{\tau_{t+1}}(\mathbf{x}^\top \tilde{\gamma}_{1,\tau_{t+1}}|\tilde{\gamma}_{\tau_{t+1}}) + \hat{g}'_{\tau_{t+1}}(\mathbf{x}^\top \tilde{\gamma}_{1,\tau_{t+1}}|\tilde{\gamma}_{\tau_{t+1}})\mathbf{X}_{i,-1}^\top(\gamma_{\tau_{t+1}} - \tilde{\gamma}_{\tau_{t+1}})\\ & \geq \hat{g}_{\tau_t}(\mathbf{x}^\top \hat{\gamma}_{1,\tau_t}|\hat{\gamma}_{\tau_t}) + \delta, \ \forall \mathbf{x} \in \mathbf{D},\end{aligned} \tag{2.11}$$

where δ is a prespecified small positive number to ensure strict inequality in (2.11). In our numerical study, we set $\delta = 10^{-4}$ as adopted by Wu and Liu (2009). Note that (2.11) can be easily solved by using the "rq.fit.fnc" function in the R package "quantreg".

Step 3: Complete down ($\gamma_{0,\tau}$).

Similar to the complete up version, we can estimate $\gamma_{0,\tau_{t-1}}$ based on $\hat{\gamma}_{\tau_t}$ by solving:

$$\begin{aligned}\hat{\gamma}_{\tau_{t-1}} &= \arg\min_{\gamma_{\tau_{t-1}}} \frac{1}{n}\sum_{i=1}^n \rho_{\tau_t}\left(Y_{i,\tau_{t-1}}^* - \hat{g}'_{\tau_{t-1}}(\mathbf{X}_i^\top \tilde{\gamma}_{1,\tau_{t-1}}|\tilde{\gamma}_{\tau_{t-1}})\mathbf{X}_{i,-1}^\top \gamma_{\tau_{t-1}}\right)\\ \text{s.t.} \quad & \hat{g}_{\tau_{t-1}}(\mathbf{x}^\top \tilde{\gamma}_{1,\tau_{t-1}}|\tilde{\gamma}_{\tau_{t-1}}) + \hat{g}'_{\tau_{t-1}}(\mathbf{x}^\top \tilde{\gamma}_{1,\tau_{t-1}}|\tilde{\gamma}_{\tau_{t-1}})\mathbf{X}_{i,-1}^\top(\gamma_{\tau_{t-1}} - \tilde{\gamma}_{\tau_{t-1}})\\ & \leq \hat{g}_{\tau_t}(\mathbf{x}^\top \hat{\gamma}_{1,\tau_t}|\hat{\gamma}_{\tau_t}) - \delta, \ \forall \mathbf{x} \in \mathbf{D}.\end{aligned}$$

Step 4: Estimate $g_{0,\tau}(\cdot)$ and $Q_{\mathbf{Y}}(\tau|\mathbf{x})$.

From the $\sqrt{nh}$-consistent initial estimator $\tilde{g}_\tau(u)$ in Step 0, and given $\hat{\gamma}_\tau$ in Steps 1-3, we estimate the link function $g_{0,\tau}(u)$ in model (1.1) at any u by

$$\begin{aligned}\hat{g}_\tau(u|\hat{\gamma}_\tau) = \ & (2\tau - 1)h_1 + \frac{\sum_{j=1}^n Y_j K_h(u - \mathbf{X}_j^\top \hat{\gamma}_{1,\tau})}{\sum_{j=1}^n K_h(u - \mathbf{X}_j^\top \hat{\gamma}_{1,\tau})}\\ & + \frac{\sum_{j=1}^n \{\tilde{g}_\tau(u) - Y_j - h_1\}K_h(u - \mathbf{X}_j^\top \hat{\gamma}_{1,\tau})I\{Y_j \leq \tilde{g}_\tau(u) - h_1\}}{\sum_{j=1}^n K_h(u - \mathbf{X}_j^\top \hat{\gamma}_{1,\tau})}\\ & + \frac{\sum_{j=1}^n \{\tilde{g}_\tau(u) - Y_j + h_1\}K_h(u - \mathbf{X}_j^\top \hat{\gamma}_{1,\tau})I\{Y_j \geq \tilde{g}_\tau(u) + h_1\}}{\sum_{j=1}^n K_h(u - \mathbf{X}_j^\top \hat{\gamma}_{1,\tau})},\end{aligned}$$

where $\hat{\gamma}_{1,\tau} = (1, \hat{\gamma}_\tau^\top)^\top$.

It is straightforward to obtain the estimator of $Q_{\mathbf{Y}}(\tau|\mathbf{x})$ as

$$\hat{Q}_{\mathbf{Y}}(\tau|\mathbf{x}) = \hat{g}_\tau(\mathbf{x}^\top \hat{\gamma}_{1,\tau}|\hat{\gamma}_\tau).$$

3. Noncrossing Single-Index CQR

3.1. *The Method*

If the model (1.1) can be rewritten as follows:

$$Q_{\mathbf{Y}}(\tau|\mathbf{x}) = g_0(\mathbf{x}^\top \gamma_{01}) + Q_\varepsilon(\tau), \tag{3.1}$$

where ε is the random error which is independent of $\mathbf{X}$, and $Q_\varepsilon(\tau)$ is the τth quantile of ε for $\tau \in (0, 1)$. The function $g_0(\cdot)$ is the unknown nonparametric smoothing functions and $\gamma_{01} = (1, \gamma_0^\top)^\top$ with $\gamma_0 \in R^{p-1}$ is the unknown parameter vector. Note that model (3.1) is a special case of model (1.1), where γ_{01} and $g_0(\cdot)$ are independent of τ. Our task is to estimate γ_0 and $g_0(\cdot)$.

For model (3.1), γ_0 is independent of τ. As discussed in Section 2.2, we can use $\bar{\gamma}_\tau$ in (2.5) to estimate γ_0 with any choice of τ. Additional efficiency gain can be achieved by combining information over multiple quantiles. We want to combine information over the K quantiles with $0 < \tau_1 < \cdots < \tau_K < 1$. Typically, we use equally spaced quantiles $\tau_k = k/(K+1)$ for $k \in \{1, \ldots, K\}$. Thus, we consider the CQR of γ_0 as follows:

$$\hat{\gamma} = \sum_{k=1}^K \nu_k \bar{\gamma}_{\tau_k}, \tag{3.2}$$

where $\sum_{k=1}^K \nu_k = 1$.

After obtaining the estimator $\hat{\gamma}$ of γ_0 in model (3.1) by (3.2), for any given point u, we consider the estimate $g_0(\cdot)$ in model (3.1). Consider weights $\mathbf{w} = (w_1, \ldots, w_K)^\top$ satisfying the constraints

$$\sum_{k=1}^K w_k = 1 \ \text{and} \ \sum_{k=1}^K w_k Q_\varepsilon(\tau_k) = 0. \tag{3.3}$$

Therefore, using (3.1) and (3.3), we have

$$\begin{aligned}\sum_{k=1}^K w_k Q_{\mathbf{Y}}(\tau_k|\mathbf{x}) &= \sum_{k=1}^K w_k \left\{g_0(\mathbf{x}^\top \gamma_{01}) + Q_\varepsilon(\tau_k)\right\}\\ &= g_0(\mathbf{x}^\top \gamma_{01})\sum_{k=1}^K w_k + \sum_{k=1}^K w_k Q_\varepsilon(\tau_k)\\ &= g_0(\mathbf{x}^\top \gamma_{01}).\end{aligned}$$

This identity suggests estimating $g_0(\cdot)$ by plugging in a consistent estimation of $Q_{\mathbf{Y}}(\tau_k|\mathbf{x})$, $k \in \{1, \ldots, K\}$. By model (1.1) and Theorem 2.3, with $\hat{\gamma}$ in (3.2), we have

$$\hat{g}(\mathbf{x}^\top \hat{\gamma}_1|\hat{\gamma}) = \sum_{k=1}^K w_k \hat{g}_{\tau_k}(\mathbf{x}^\top \hat{\gamma}_1|\hat{\gamma}), \tag{3.4}$$

where $\hat{g}_{\tau_k}(\cdot|\gamma)$ is defined in (2.3) and $\hat{\gamma}_1 = (1, \hat{\gamma}^\top)^\top$. Moreover, for any given point u,

$$\hat{g}(u|\hat{\gamma}) = \sum_{k=1}^K w_k \hat{g}_{\tau_k}(u|\hat{\gamma}). \quad (3.5)$$

From (3.2), (3.4), and (3.5), it is easy to see that NCQR estimators with τ_k in Section 2 are a special case of noncrossing composite quantile regression (NCCQR) estimators when the kth position of vector $\mathbf{v}$ or $\mathbf{w}$ is 1 and the rest is 0. Therefore, under the choice of appropriate weight $\mathbf{v}$ or $\mathbf{w}$, the estimated efficiency of NCCQR is better than that of NCQR.

3.2. Asymptotic Normality of the Estimators

To establish the asymptotic properties of the proposed estimators, the following technical conditions are imposed.

Condition 7: The density function of $\mathbf{X}^\top\gamma_1$ is positive and uniformly continuous for γ_1 in a neighborhood of γ_{01}. Furthermore, the density of $\mathbf{X}^\top\gamma_{01}$ is continuous and bounded away from 0 and ∞ on its support.

Condition 8: The function $g_0(\cdot)$ has a continuous and bounded second derivative.

Condition 9: $f_\varepsilon(\cdot)$ is the density function of ε. For any $\tau \in (0, 1)$, $f_\varepsilon(Q_\varepsilon(\tau))$ is uniformly bounded away from zero and infinity.

Remark 3.1. Conditions 7 and 8 are similar to Conditions 1 and 2 under different models (1.1) and (3.1). Condition 9 is commonly assumed to derive the asymptotic normality of the quantile regression estimator; see Koenker (2005).

3.2.1. Asymptotic Normality of Parametric Regression Estimators

Theorem 3.1. Under regular Conditions 3-9, assume that $\sum_{k=1}^K v_k = 1$; if $n \to \infty$, $h_1 \to 0$, $nh^9 \to 0$, $nh^5h_1^4 \to 0$, and $nh \to \infty$, then,

$$\sqrt{n}(\hat{\gamma} - \gamma_0) \xrightarrow{L} \mathcal{N}\left(0, R_1(\mathbf{v})\mathbf{S}_2^{-1}\right),$$

where $R_1(\mathbf{v}) = \sum_{k=1}^K \sum_{k'=1}^K v_k v_{k'} \{\min(\tau_k, \tau_{k'}) - \tau_k\tau_{k'}\}/\{f_\varepsilon(Q_\varepsilon(\tau_k)) f_\varepsilon(Q_\varepsilon(\tau_{k'}))\}$, $\mathbf{v} = (v_1, \ldots, v_K)^\top$, and $\mathbf{S}_2 = E\left[\{g_0'(\mathbf{X}^\top\gamma_{01})\}^2 \tilde{\mathbf{X}}_{-1}\tilde{\mathbf{X}}_{-1}^\top\right]$ is assumed to be a nonsingular matrix.

From Theorem 3.1, we find that the asymptotic variance of $\hat{\gamma}$ depends on $\mathbf{v}$ only through $R_1(\mathbf{v})$. Thus, the optimal choice of weights for maximizing the efficiency of the estimator $\hat{\gamma}$ is

$$\begin{aligned}\mathbf{v}_{\text{opt}} &= \arg\min_{\mathbf{v}} R_1(\mathbf{v}), \text{ such that } \sum_{k=1}^K v_k = 1 \\ &= \mathbf{D}\mathbf{T}^{-1}\mathbf{D}\mathbf{1}/(\mathbf{1}^\top\mathbf{D}\mathbf{T}^{-1}\mathbf{D}\mathbf{1}),\end{aligned} \quad (3.6)$$

where $\mathbf{D} = \text{diag}\{f_\varepsilon(Q_\varepsilon(\tau_1)), \ldots, f_\varepsilon(Q_\varepsilon(\tau_K))\}$, $\mathbf{T}$ is the $K \times K$ matrix with entries $\min(\tau_k, \tau_{k'}) - \tau_k\tau_{k'}$, $k, k' \in \{1, \ldots, K\}$, and $\mathbf{1}$ is the $K \times 1$ matrix with entries 1. Thus, the optimal asymptotic variance using these weights (3.6) is $(\mathbf{f}^\top\mathbf{T}^{-1}\mathbf{f})^{-1}\mathbf{S}_2^{-1}$, where $\mathbf{f} = (f_\varepsilon(Q_\varepsilon(\tau_1)), \ldots, f_\varepsilon(Q_\varepsilon(\tau_K)))^\top$. Note that the optimal asymptotic variance is the same as the weighted CQR proposed by Jiang et al. (2016). Therefore, our proposed estimator is more efficient than many estimation methods. The details can be seen in Section 2.4 in Jiang et al. (2016).

It can be seen from (3.6) that the optimal weight vector $\mathbf{v}_{\text{opt}}$ involves unknown parameters $f_\varepsilon(\cdot)$ and $Q_\varepsilon(\tau_k)$, $k = 1, \ldots, K$. We can obtain the estimator of $\mathbf{v}_{\text{opt}}$ as

$$\hat{\mathbf{v}} = \hat{\mathbf{D}}\mathbf{T}^{-1}\hat{\mathbf{D}}\mathbf{1}/(\mathbf{1}^\top\hat{\mathbf{D}}\mathbf{T}^{-1}\mathbf{D}\mathbf{1}),$$

where $\hat{\mathbf{D}} = \text{diag}\{\hat{f}_{\hat{\varepsilon}}(\hat{Q}_{\hat{\varepsilon}}(\tau_1)), \ldots, \hat{f}_{\hat{\varepsilon}}(\hat{Q}_{\hat{\varepsilon}}(\tau_K))\}$. $\hat{Q}_{\hat{\varepsilon}}(\tau)$ is the sample τ-quantile of $\hat{\varepsilon} = \mathbf{Y} - \hat{g}(\mathbf{X}^\top\hat{\gamma}_1)$, where $\hat{\gamma}_1$ and $\hat{g}(\cdot)$ are obtained by (2.4) and (2.3) with $\tau = 0.5$. $\hat{f}_{\hat{\varepsilon}}(u) = n^{-1}\sum_{i=1}^n K_{h_2}(\hat{\varepsilon}_i - u)$ and h_2 can be chosen as $h_2 = 0.9 \times \min\{\text{std}(\hat{\varepsilon}), \text{IQR}(\hat{\varepsilon})/1.34\} \times n^{-1/5}$, where std and IQR denote the sample standard deviation and sample interquantile, respectively (Silverman 1986).

3.2.2. Asymptotic Normality of the Nonparametric Functions

Theorem 3.2. Under the same conditions as in Theorem 3.1, assume that $\sum_{k=1}^K w_k = 1$ and $\sum_{k=1}^K w_k Q_\varepsilon(\tau_k) = 0$; then, for an interior point u,

$$\sqrt{nh}\left\{\hat{g}(u|\hat{\gamma}) - g_0(u) - \alpha(u)\right\} \xrightarrow{L} \mathcal{N}\left(0, \frac{\nu_0 R_2(\mathbf{w})}{f_{\mathbf{U}_0}(u)}\right),$$

where $f_{\mathbf{U}_0}(\cdot)$ is the marginal density function of $\mathbf{U}_0 = \mathbf{X}^\top\gamma_{01}$, and

$$\alpha(u) = f^{-1}(g_0(u)|u)\left\{\tfrac{1}{2}h^2\mu_2\beta(u) + \tfrac{1}{6}h_1^2 \tfrac{\partial^2 F(y|u)}{\partial y^2}\Big|_{y=g_0(u)}\right\},$$

$$\beta(u) = \tfrac{\partial^2 F(g_0(u)|u)}{\partial u^2} + 2f_{\mathbf{U}_0}^{-1}(u)f_{\mathbf{U}_0}'(u)\tfrac{\partial F(g_0(u)|u)}{\partial u},$$

$$R_2(\mathbf{w}) = \sum_{k=1}^K \sum_{k'=1}^K w_w w_{k'} \frac{\min(\tau_k, \tau_{k'}) - \tau_k\tau_{k'} - 1/3h_1 f_\varepsilon(Q_\varepsilon(\min(\tau_k, \tau_{k'})))}{f_\varepsilon(Q_\varepsilon(\tau_k))f_\varepsilon(Q_\varepsilon(\tau_{k'}))}.$$

Furthermore, when $nh^4 \to \infty$, for any interior point $\mathbf{x}$,

$$\sqrt{nh}\left\{\hat{g}(\mathbf{x}^\top\hat{\gamma}|\hat{\gamma}) - g_0(\mathbf{x}^\top\gamma_0) - \alpha(\mathbf{x}^\top\gamma_0)\right\} \xrightarrow{L} \mathcal{N}\left(0, \frac{\nu_0 R_2(\mathbf{w})}{f_{\mathbf{U}_0}(\mathbf{x}^\top\gamma_0)}\right).$$

It is easy to see that the bias of $\hat{g}(u|\hat{\gamma})$ or of $\hat{g}(\mathbf{x}^\top\hat{\gamma}|\hat{\gamma})$ is free of the choice of weight vector $\mathbf{w}$, and only the variance term $R_2(\mathbf{w})$ depends on the weight vector $\mathbf{w}$. Then, the optimal weights correspond to the minimum $R_2(\mathbf{w})$,

$$\begin{aligned}\mathbf{w}_{\text{opt}} &= \arg\min_{\mathbf{w}} R_2(\mathbf{w}), \text{ such that } \sum_{k=1}^K w_k = 1 \\ &\quad \text{and } \sum_{k=1}^K w_k Q_\varepsilon(\tau_k) = 0 \\ &= \frac{(\mathbf{r}^\top\Omega^{-1}\mathbf{r})\Omega^{-1}\mathbf{1} - (\mathbf{1}^\top\Omega^{-1}\mathbf{r})\Omega^{-1}\mathbf{r}}{(\mathbf{r}^\top\Omega^{-1}\mathbf{r})(\mathbf{1}^\top\Omega^{-1}\mathbf{1}) - (\mathbf{1}^\top\Omega^{-1}\mathbf{r})^2},\end{aligned} \quad (3.7)$$

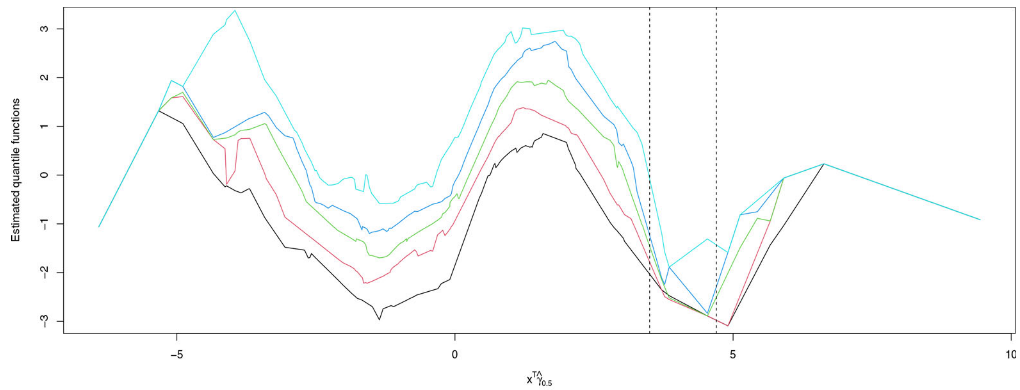
where $\mathbf{r}$ is a K-dimensional column vector with kth element $Q_\varepsilon(\tau_k)$ and Ω is the $K \times K$ matrix with the (k, k')th element

$$\{\min(\tau_k, \tau_{k'}) - \tau_k\tau_{k'} - 1/3h_1 f_\varepsilon(Q_\varepsilon(\min(\tau_k, \tau_{k'}))\} / \{f_\varepsilon(Q_\varepsilon(\tau_k))f_\varepsilon(Q_\varepsilon(\tau_{k'}))\}.$$

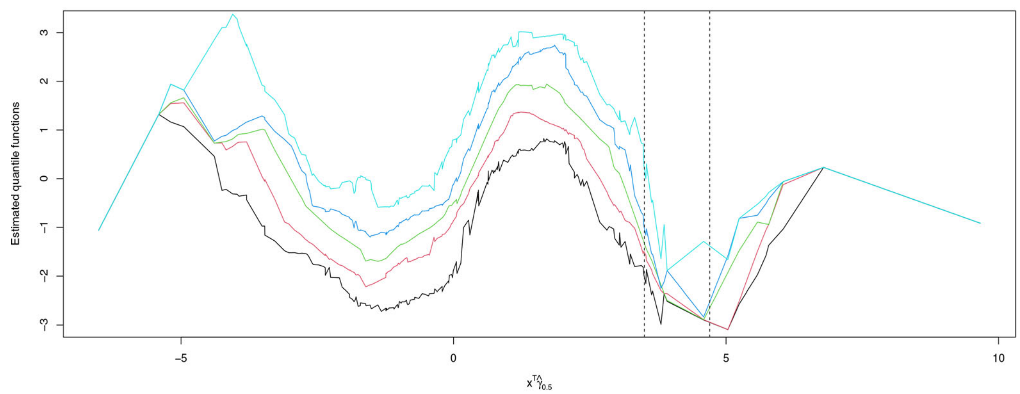
Note that $\mathbf{w}_{\text{opt}}$ also involves unknown parameters $f_\varepsilon(\cdot)$ and $\{Q_\varepsilon(\tau_k)\}_{k=1}^K$ as $\mathbf{v}_{\text{opt}}$. Therefore, we can use a similar method for $\mathbf{v}_{\text{opt}}$ to obtain the estimator of $\mathbf{w}_{\text{opt}}$.

4. Numerical Studies

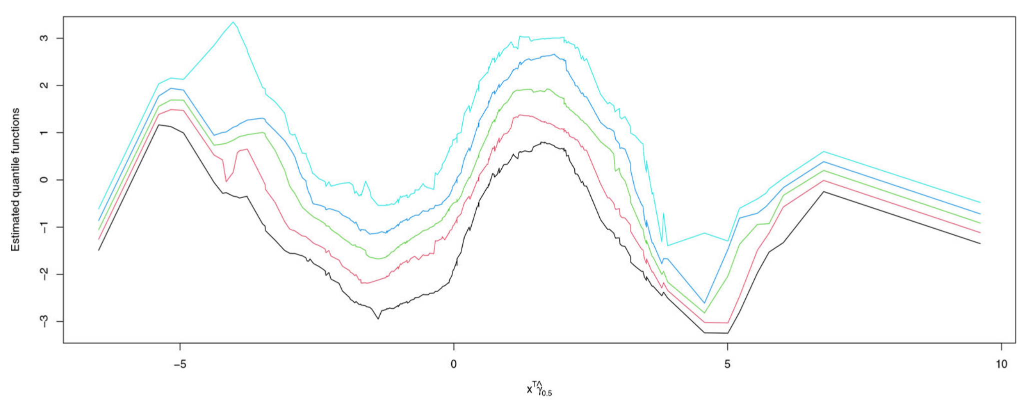
In this section, we first use Monte Carlo simulation studies to assess the finite sample performance of the proposed procedures. We then demonstrate the application of the proposed



(a) NIQR-W



(b) NIQR-C



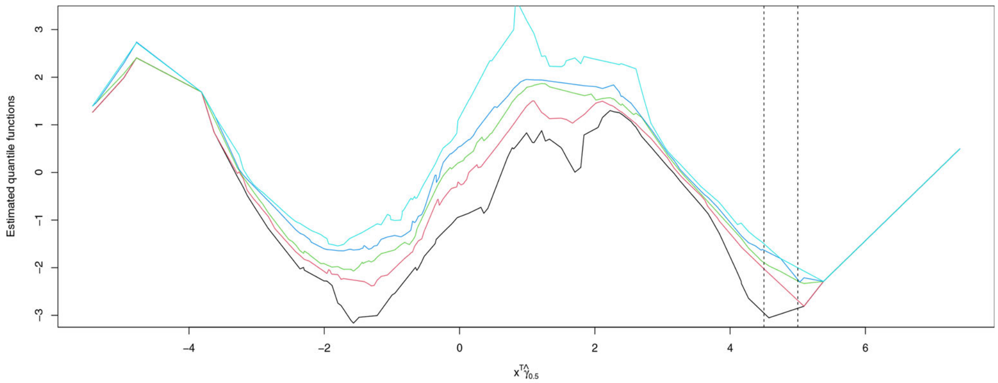
(c) NCQR

Figure 1. The estimated conditional quantile functions for model (4.1) with Case 1 under different quantiles and estimation methods. The abscissa is $\mathbf{x}^{\top}\hat{\gamma}_{0.5}$, where $\hat{\gamma}_{0.5}$ is obtained by NIQR-W, NIQR-C, or NCQR at $\tau = 0.5$.

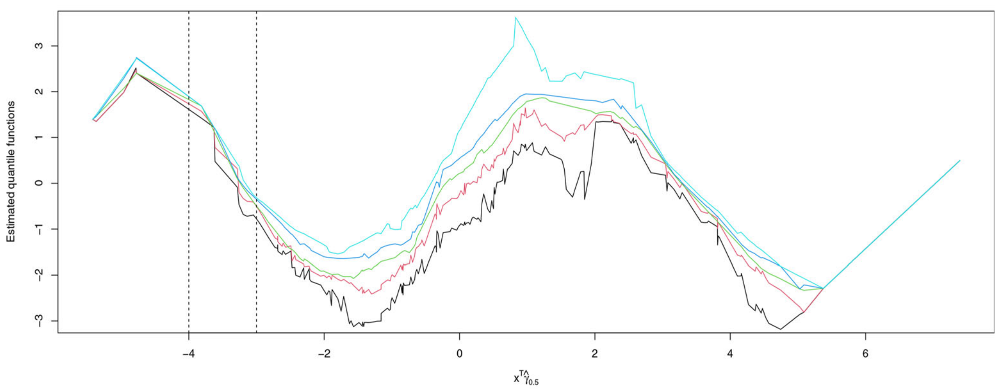
methods with a real data analysis. All programs are written in R code. The Gaussian kernel $K(u) = (\sqrt{2\pi})^{-1}\exp(-u^2/2)$ is used in this section.

4.1. Simulation Example 1

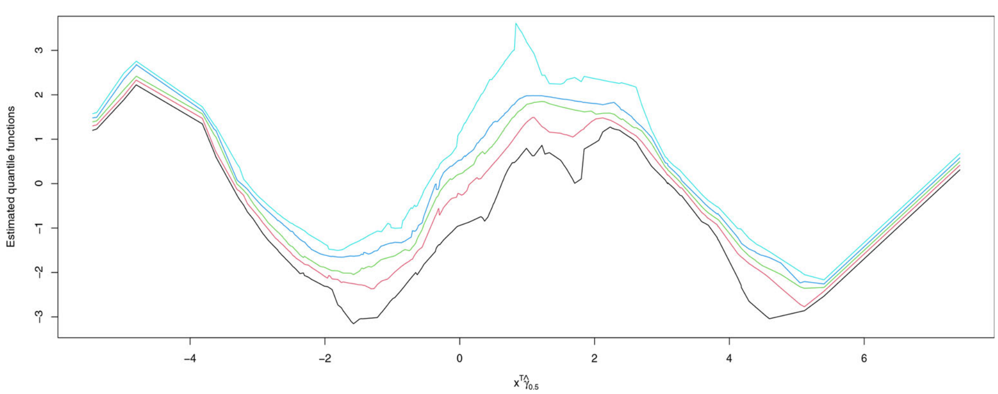
We use simulation studies to illustrate the improvement of our NCQR for the single-index model in Section 2 by comparing it



(a) NIQR-W



(b) NIQR-C



(c) NCQR

Figure 2. The estimated conditional quantile functions for model (4.1) with Case 2 under different quantiles and estimation methods. The abscissa is $\mathbf{x}^\top \hat{\gamma}_{0.5}$, where $\hat{\gamma}_{0.5}$ is obtained by NIQR-W, NIQR-C, or NCQR at $\tau = 0.5$.

to the naive individual quantile regression (NIQR). We consider two NIQR methods: NIQR-W (Wu et al. 2010) and NIQR-C (Christou and Akritas 2016). We use the following model (4.1) to demonstrate that NIQR estimates may suffer from quantile crossing, while our proposed NCQR method can avoid it.

$$\mathbf{Y} = 2\sin(\mathbf{X}^\top \gamma_{01}) + \sigma(\mathbf{X})\varepsilon, \tag{4.1}$$

where $\mathbf{X}_j \sim N(0,1)$ for $j = 1,2$, $\gamma_{01} = (1,2)^\top$, and the sample size is fixed at $n = 200$. Two error distributions of ε are considered: a standard normal distribution (N(0,1)) and a t distribution with 3 degrees of freedom (t(3)). We consider the following two cases:

Case 1 (Normal errors): $\sigma(\mathbf{X}) = 1$ and $\varepsilon \sim N(0,1)$, and

Case 2 (Heteroscedastic errors): $\sigma(\mathbf{X}) = 0.5\sqrt{1+\cos(\mathbf{X}^\top\gamma_{01})}$ and $\varepsilon \sim t(3)$.

Quantile functions are estimated at $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$. The NIQR-W, NIQR-C and NCQR estimated quantile functions are plotted in Figures 1 and 2 for different cases and quantiles. We can see that NIQR-W and NIQR-C suffer from quantile crossing (see the area between the dotted lines). However, by enforcing our noncrossing adjustment in Section 2, our new estimates (NCQR) do not cross each other.

4.2. Simulation Example 2

To further compare with the NIQR, we consider the following four simulation designs:

Design 1 (Homoscedastic model): $\mathbf{Y} = \exp\left(\mathbf{X}^\top\gamma_{01}\right)+\varepsilon$, where $\mathbf{X} \sim U(0,1)$, $\gamma_{01} = (1,2)^\top$, $\varepsilon \sim N(0,1)$, and $n = 200$.

Design 2 (Heteroscedastic model): $\mathbf{Y} = 5\cos\left(2\pi\mathbf{X}^\top\gamma_0\right) + \sqrt{1+\left(\mathbf{X}^\top\gamma_{01}\right)^2}\ \varepsilon$, where $\mathbf{X} \sim U(0,1)$, $\gamma_{01} = (1,-1)^\top$, $\varepsilon \sim N(0,1)$, and $n = 200$.

Design 3 (High-dimensional homoscedastic model): $\mathbf{Y} = \left(\mathbf{X}^\top\gamma_{01}\right)^2 + \varepsilon$, where $\mathbf{X}$ is drawn from a multivariate uniform distribution on the $[0,1]^p$ with covariance matrix $\Sigma_{ij} = 0.5^{i-j}$ for $1 \le i,j \le p$, $p = 100$, $\gamma_{01} = (1,2,1,0,\dots,0)^\top$, $\varepsilon \sim t(5)$, and $n = 1000$.

Design 4 (High-dimensional heteroscedastic model): $\mathbf{Y} = \left(\mathbf{X}^\top\gamma_{01}\right)^2 + \sqrt{1+sin\left(\mathbf{X}^\top\gamma_{01}\right)}\ \varepsilon$, where $\mathbf{X}$ is drawn from a multivariate uniform distribution on the $[0,1]^p$ with covariance matrix $\Sigma_{ij} = 0.5^{i-j}$ for $1 \le i,j \le p$, $p = 100$, $\gamma_{01} = (1,2,1,0,\dots,0)^\top$, $\varepsilon \sim N(0,1)$, and $n = 1000$.

Quantiles $\tau = 0.1,\dots,0.9$ are fitted to the data. We simulate 500 replicates, and all samples present crossing issues when using NIQR. To compare the two methods, the empirical root mean integrated squared errors (RMISE) are computed for each design and τ.

$$\mathrm{RMISE} = \left[\frac{1}{n}\sum_{i=1}^n\left\{\hat{Q}_{\mathbf{Y}}(\tau|\mathbf{X}_i) - Q_{\mathbf{Y}}(\tau|\mathbf{X}_i)\right\}^2\right]^{1/2},$$

where $\hat{Q}_{\mathbf{Y}}(\tau|\cdot)$ is the estimated function of the true function $Q_{\mathbf{Y}}(\tau|\cdot)$. Tables 1–4 present the average RMISE over the 500 datasets along with its estimated standard error for four designs and three methods, respectively. In each of the settings considered, the proposed NCQR estimators of quantile functions give significantly better estimates for all qantiles based on their smaller RMISEs. Since the true curves do not cross, it is expected that the NCQR estimator of the quantile function performs better.

Table 1. This table compares NIQR (NIQR-W and NIQR-C) and NCQR based on the mean values and standard errors (in parentheses) of RMISE based on 500 estimates for Design 1 under different quantiles.

τ	NIQR-W	NIQR-C	NCQR
0.1	0.386 (0.103)	0.389 (0.101)	0.352 (0.084)
0.2	0.339 (0.077)	0.343 (0.078)	0.317 (0.069)
0.3	0.315 (0.073)	0.323 (0.087)	0.297 (0.066)
0.4	0.305 (0.069)	0.310 (0.070)	0.289 (0.064)
0.5	0.298 (0.066)	0.304 (0.073)	0.285 (0.060)
0.6	0.303 (0.069)	0.310 (0.072)	0.290 (0.062)
0.7	0.316 (0.079)	0.311 (0.072)	0.300 (0.067)
0.8	0.332 (0.077)	0.340 (0.080)	0.315 (0.067)
0.9	0.391 (0.098)	0.399 (0.102)	0.366 (0.085)

Table 2. This table compares NIQR (NIQR-W and NIQR-C) and NCQR based on the mean values and standard errors (in parentheses) of RMISE based on 500 estimates for Design 2 under different quantiles.

τ	NIQR-W	NIQR-C	NCQR
0.1	0.608 (0.118)	0.598 (0.111)	0.532 (0.091)
0.2	0.520 (0.102)	0.512 (0.100)	0.471 (0.083)
0.3	0.482 (0.092)	0.473 (0.093)	0.441 (0.077)
0.4	0.471 (0.081)	0.457 (0.085)	0.433 (0.072)
0.5	0.469 (0.081)	0.451 (0.075)	0.431 (0.071)
0.6	0.476 (0.088)	0.458 (0.085)	0.439 (0.077)
0.7	0.491 (0.095)	0.473 (0.097)	0.453 (0.083)
0.8	0.530 (0.100)	0.510 (0.097)	0.484 (0.084)
0.9	0.602 (0.117)	0.592 (0.112)	0.538 (0.097)

Table 3. This table compares NIQR (NIQR-W and NIQR-C) and NCQR based on the mean values and standard errors (in parentheses) of RMISE based on 500 estimates for Design 3 under different quantiles.

τ	NIQR-W	NIQR-C	NCQR
0.1	0.702 (0.064)	0.967 (0.065)	0.666 (0.065)
0.2	0.569 (0.032)	0.714 (0.048)	0.544 (0.029)
0.3	0.496 (0.036)	0.631 (0.029)	0.479 (0.034)
0.4	0.475 (0.048)	0.591 (0.032)	0.459 (0.050)
0.5	0.457 (0.050)	0.576 (0.050)	0.446 (0.052)
0.6	0.475 (0.040)	0.607 (0.030)	0.456 (0.042)
0.7	0.527 (0.050)	0.624 (0.062)	0.496 (0.050)
0.8	0.570 (0.057)	0.736 (0.057)	0.543 (0.056)
0.9	0.696 (0.055)	0.886 (0.047)	0.663 (0.053)

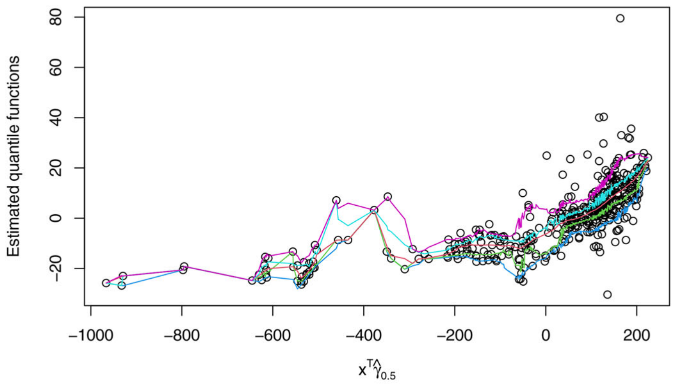
Table 4. This table compares NIQR (NIQR-W and NIQR-C) and NCQR based on the mean values and standard errors (in parentheses) of RMISE based on 500 estimates for Design 4 under different quantiles.

τ	NIQR-W	NIQR-C	NCQR
0.1	0.542 (0.022)	0.565 (0.044)	0.511 (0.020)
0.2	0.424 (0.026)	0.452 (0.022)	0.396 (0.024)
0.3	0.358 (0.025)	0.393 (0.027)	0.334 (0.025)
0.4	0.331 (0.023)	0.353 (0.032)	0.308 (0.023)
0.5	0.315 (0.018)	0.326 (0.030)	0.303 (0.017)
0.6	0.321 (0.023)	0.335 (0.027)	0.308 (0.022)
0.7	0.341 (0.018)	0.350 (0.023)	0.320 (0.028)
0.8	0.389 (0.027)	0.402 (0.029)	0.365 (0.026)
0.9	0.523 (0.048)	0.535 (0.047)	0.506 (0.046)

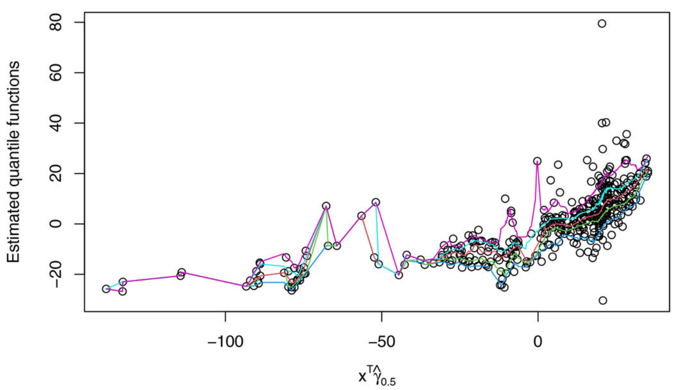
4.3. Simulation Example 3

We use Monte Carlo simulation studies to assess the finite sample performance of the proposed NCCQR method for single-index models in Section 3. By Tables 1 and 2 in Jiang et al. (2016), $K = 9$ is a good choice for NCCQR. Therefore, we only consider $K = 9$ in this section (NCCQR$_9$). Furthermore, we include four competitors: (i) the NIQR-W with $\tau = 0.5$ (NIQR-W$_{0.5}$); (ii) the NIQR-C with $\tau = 0.5$ (NIQR-C$_{0.5}$); (iii) the NCQR with $\tau = 0.5$ (NCQR$_{0.5}$); and (iv) the weighted CQR with $K = 9$ (WCQR$_9$) proposed by Jiang et al. (2016). For the sake of comparison, we modify WCQR$_9$ by setting its first component equal to 1. The data are generated according to the following models:

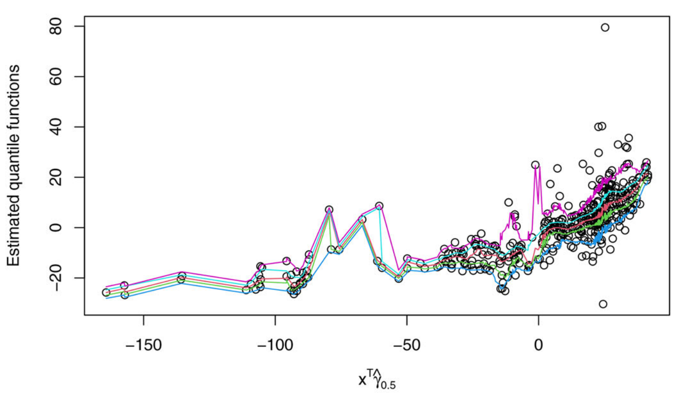
Design 5: $\mathbf{Y} = 2\sin\left(\mathbf{X}^\top\gamma_0\right) + \varepsilon$, where $\mathbf{X} \sim N(0,1)$, $\gamma_0 = (1,2)^\top$, $\varepsilon \sim N(0,1)$, and $n = 200$.



(a) NIQR-W



(b) NIQR-C



(c) NCQR

Figure 3. The estimated conditional quantile functions for real estate valuation data under different quantiles and estimation methods. The abscissa is $\mathbf{x}^\top \hat{\gamma}_{0.5}$, where $\hat{\gamma}_{0.5}$ is obtained by NIQR-W, NIQR-C, or NCQR at $\tau = 0.5$.

Table 5. This table compares five estimation methods (NIQR-$W_{0.5}$; NIQR-$C_{0.5}$; $NCQR_{0.5}$; $WCQR_9$; $NCCQR_9$) based on the mean values and standard errors (in parentheses) of AAB and AIAB based on 500 estimates for Example 3 under different designs.

Model	Result	NIQR-$W_{0.5}$	NIQR-$C_{0.5}$	$NCQR_{0.5}$	$WCQR_9$	$NCCQR_9$
Design 5	AAB	0.162 (0.129)	0.176 (0.138)	0.136 (0.102)	0.133 (0.108)	0.122 (0.109)
	AIAB	0.263 (0.054)	0.257 (0.055)	0.247 (0.047)	0.231 (0.050)	0.224 (0.049)
Design 6	AAB	0.039 (0.017)	0.043 (0.022)	0.035 (0.020)	0.035 (0.016)	0.033 (0.018)
	AIAB	0.262 (0.039)	0.280 (0.049)	0.254 (0.035)	0.251 (0.038)	0.241 (0.037)
Design 7	AAB	0.090 (0.006)	0.099 (0.004)	0.085 (0.006)	0.085 (0.005)	0.080 (0.005)
	AIAB	0.317 (0.016)	0.335 (0.015)	0.312 (0.017)	0.292 (0.017)	0.287 (0.019)

Table 6. This table presents the single-index coefficient estimates for real estate valuation data under different estimation methods and quantiles.

Method	τ	Number	Meter	Year	Date
NIQR-W	0.1	1.000	−0.024	−0.307	5.361
	0.3	1.000	−0.024	−0.307	5.382
	0.5	1.000	−0.024	−0.315	5.557
	0.7	1.000	−0.026	−0.352	6.189
	0.9	1.000	−0.025	−0.321	5.979
NIQR-C	0.1	1.000	−0.063	−0.796	13.587
	0.3	1.000	−0.111	−1.585	19.364
	0.5	1.000	−0.175	−2.119	36.924
	0.7	1.000	−0.216	−3.407	60.921
	0.9	1.000	−0.170	−2.463	42.316
NCQR	0.1	1.000	−0.023	−0.340	5.651
	0.3	1.000	−0.024	−0.336	5.624
	0.5	1.000	−0.029	−0.421	6.974
	0.7	1.000	−0.030	−0.509	7.798
	0.9	1.000	−0.028	−0.407	8.326
WCQR9		1.000	−0.025	−0.322	5.607
NCCQR9		1.000	−0.026	−0.395	6.489

Design 6: $\mathbf{Y} = \left(\mathbf{X}^\top \gamma_0\right)^2 + \varepsilon$, where $\mathbf{X} \sim U(1,2)$, $\gamma_0 = (1,1,1)^\top$, $\varepsilon \sim t(5)$, and $n = 200$.

Design 7: $\mathbf{Y} = \exp\left(\mathbf{X}^\top \gamma_0\right) + \varepsilon$, where $\mathbf{X}$ is drawn from a multivariate uniform distribution on the $[0,1]^p$ with $\Sigma_{ij} = 0.5^{i-j}$ for $1 \le i,j \le p$, $p = 100$, $\gamma_0 = (1,1,-1,0,\ldots,0)^\top$, $\varepsilon \sim N(0,1)$, and $n = 1000$.

To compare the five methods, the average absolute bias (AAB) of $\hat{\gamma}$ and the average integrated absolute bias (AIAB) of $\hat{g}(\cdot)$ are computed for each design.

$$AAB = \frac{1}{p-1} \sum_{j=2}^{p} |\hat{\gamma}_j - \gamma_{0,j}|,$$
$$AIAB = \frac{1}{n_{\text{grid}}} \sum_{i=1}^{n_{\text{grid}}} \left|\hat{g}(u_i) - g_0(u_i)\right|,$$

where $u_i, i = 1, \ldots, n_{\text{grid}}$ are grid points of the support of $\mathbf{X}^\top \gamma_0$, and $n_{\text{grid}} = n$. The simulation results are summarized in Table 5. From the Table 5, we can see that NCCQR$_9$ performs well under different models and high-dimensional data. Moreover, in each of the designs considered, the proposed NCQR$_{0.5}$ estimators of the parametric and nonparametric parts give significantly better results than those of NIQR$_{0.5}$ (NIQR-W$_{0.5}$ and NIQR-C$_{0.5}$) based on their smaller AAB and AIAB. Since the true curves do not cross, it is expected that the NCQR estimator of the nonparametric part performs better. By using the noncrossing estimator of the nonparametric part, the proposed NCQR estimators of the parametric part improve estimation efficiency.

4.4. Data Example: Real Estate Valuation Data

As an illustration, we now apply the proposed methodology to real estate valuation data. The dataset is collected from the public database of the Ministry of the Interior during the period of June 2012 to May 2013 from two districts in Taipei City and two districts in New Taipei City (*https://archive.ics.uci.edu/~ml/datasets/Real+estate+valuation+data+set*). The data contain 414 observations on six variables, and the dependent variable of interest is the residential housing price per unit area (Price). Based on a related research Yeh and Hsu (2018), four appraisal factors (independent variables) are chosen: the distance to the nearest MRT station (Meter); the number of convenience stores in the living circle on foot (Number); house age (Year); and transaction date (Date). All variables are centered around zero. The following single-index quantile regression is used to fit the data:

$$\begin{aligned} & Q_{\text{Price}}(\tau|\text{Number};\ \text{Meter};\ \text{Year};\ \text{Date}) \\ = \ & g_\tau\{\gamma_{1,\tau}\text{Number} + \gamma_{2,\tau}\text{Meter} + \gamma_{3,\tau}\text{Year} + \gamma_{4,\tau}\text{Date}\}. \end{aligned}$$

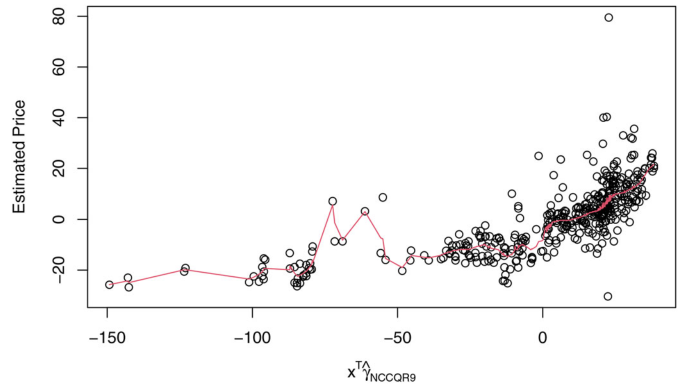


Figure 4. The estimated Price by NCCQR$_9$ for real estate valuation data. The dots are the observations Price and the curve is the estimated Price by the proposed NCCQR$_9$ method. The abscissa is $\mathbf{x}^\top \hat{\gamma}_{\text{NCCQR}_9}$.

Table 7. The ASPE of five estimation methods (NIQR-W$_{0.5}$; NIQR-C$_{0.5}$; NCQR$_{0.5}$; WCQR$_9$; NCCQR$_9$) for real estate valuation data.

Method	NIQR-W$_{0.5}$	NIQR-C$_{0.5}$	NCQR$_{0.5}$	WCQR$_9$	NCCQR$_9$
ASPE	42.617	44.204	42.027	40.857	40.805

The estimated coefficients under different methods for the above model are summarized in Table 6. Figure 3 shows the estimated $Q_{\text{Price}}(\tau|\text{Number};\ \text{Meter};\ \text{Year};\ \text{Date})$ of quantiles $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$ along with the data by NIQR-W, NIQR-C, and NCQR. There are 10 data crossings for NIQR-W and 10 data crossings for NIQR-C. In contrast, the NCQR improves the estimates by removing the crossing. Figure 4 shows the estimated Price by the NCCQR$_9$ method along with the observations. This suggests that the estimated Price by NCCQR$_9$ provides a good fit to the data.

Furthermore, to compare the performance of the five methods (NIQR-W$_{0.5}$; NIQR-C$_{0.5}$; NCQR$_{0.5}$; WCQR$_9$; NCCQR$_9$), we evaluate the average square prediction error (ASPE). The first 320 data points are used for the estimation, and the remaining 94 data points are used for the prediction. Therefore,

$$\text{ASPE} = \frac{1}{\tilde{n}} \sum_{i=1}^{\tilde{n}} (\text{Price}_i - \widehat{\text{Price}}_i)^2,$$

where $\widehat{\text{Price}}$ is the fitted value of Price and $\tilde{n} = 94$. The results of ASPEs under different methods are presented in Table 7. It is easy to see that the performances of NCQR$_{0.5}$ are better than those of NIQR-W$_{0.5}$ and NIQR-C$_{0.5}$. The results of NCCQR$_9$ are the best based on the smallest ASPE.

5. Conclusion

In this article, we have considered the estimation problem of the single-index quantile regression without quantile crossing. An effective and simple-kernel estimation method of the nonparametric part in model (1.1) has been provided for both standard single-index quantile regression and single-index CQR.

Our proposed method gives a closed form expression for the estimator of the nonparametric part and ensures noncrossing. Thus, the restrictive estimators of unknown parameters are valid, which guarantees the noncrossing estimator of $Q_{\mathbf{Y}}(\tau|\mathbf{x})$. Simulations have demonstrated that the proposed noncrossing method not only helps to provide more meaningful results, but also improves the estimation accuracy of the resulting regression functions. The real data application example further highlights its value in applied settings.

Supplementary Material

The proofs of the proposed theorems are given in the supplementary material file.

Acknowledgments

We thank the editor, the associate editor and the two reviewers for their considered and detailed comments on our article, which have helped us to improve it.

Funding

This work was supported by the National Social Science Foundation of China (Series number: 21BTJ040) and the National Natural Science Foundation of China (Series number: 11801069).

References

Andriyana, Y., Gijbels, I., and Verhasselt, A. (2018), "Quantile Regression in Varying-Coefficient Models: Non-Crossing Quantile Curves and Heteroscedasticity," *Statistical Papers*, 59, 1589–1621. [310]

Bondell, H., Reich, B., and Wang, H. (2010), "Non-Crossing Quantile Regression Curve Estimation," *Biometrika*, 97, 825–838. [310]

Chaudhuri, P. (1991), "Nonparametric Estimates of Regression Quantiles and Their Local Bahadur Representation," *Annals of Statistics*, 2, 760–777. [309]

Chaudhuri, P., Doksum, K., and Samarov, A. (1997), "On Average Derivative Quantile Regression," *Annals of Statistics*, 25, 715–744. [309]

Chen, D., Shujie, M., Liping, Z., and Feng, X. (2021), "Estimation and Inference for Non-Crossing Multiple-Index Quantile Regression," *Scientia Sinica Mathematica*, 51, 631–658. [310]

Chernozhukov, V., Fernandez-Val, I., and Galichon, A. (2010), "Quantile and Probability Curves Without Crossing," *Econometrica*, 78, 1093–1125. [310]

Christou, E., and Akritas, M. (2016), "Single Index Quantile Regression for Heteroscedastic Data," *Journal of Multivariate Analysis*, 150, 169–182. [309,310,311,312,313,316]

Dette, H., and Volgushev, S. (2008), "Non-Crossing Nonparametric Estimates of Quantile Curves," *Journal of the Royal Statistical Society*, Series B, 70, 609–627. [310]

Guerre, E., and Sabbah, C. (2012), "Uniform Bias Study and Bahadur Representation for Local Polynomial Estimators of the Conditional Quantile Function," *Econometric Theory*, 28, 87–129. [309]

Hall, P., Wolff, R., and Yao, Q. (1999), "Methods for Estimating a Conditional Distribution Function," *Journal of the American Statistical Association*, 94, 154–163. [310]

He, X. (1997), "Quantile Curves Without Crossing," *The American Statistician*, 51, 186–192. [310]

Huang, M., and Nguyen, C. (2018), "A Nonparametric Approach for Quantile Regression," *Journal of Statistical Distributions and Applications*, 5, 3. [309]

Ichimura, H. (1993), "Semiparametric Least Squares (sls) and Weighted sls Estimation of Single Index Models," *Journal of Econometrics*, 58, 71–120. [309]

Jiang, R., Qian, W.-M., and Zhou, Z.-G. (2016), "Weighted Composite Quantile Regression for Single-Index Models," *Journal of Multivariate Analysis*, 148, 34–48. [309,310,314,317]

Jiang, R., and Yu, K. (2020), "Single-Index Composite Quantile Regression for Massive Data," *Journal of Multivariate Analysis*, 180, 104669. [310]

Kim, J., Cho, H., and Bang, S. (2019), "Unified Non-Crossing Multiple Quantile Regressions Tree," *Journal of Computational and Graphical Statistics*, 28, 454–465. [310]

Koenker, R. (2005), *Quantile Regression*, New York: Cambridge University Press. [309,314]

Kong, E., Linton, O., and Xia, Y. (2013), "Global Bahadur Representation for Nonparametric Censored Regression Quantiles and Its Applications," *Econometric Theory*, 29, 941–968. [309]

Kong, E., and Xia, Y. (2012), "A Single-Index Quantile Regression Model and Its Estimation," *Econometric Theory*, 28, 730–768. [309,311,313]

Li, Q., Lin, J., and Racine, J. (2013), "Optimal Bandwidth Selection for Nonparametric Conditional Distribution and Quantile Functions," *Journal of Business & Economic Statistics*, 31, 57–65. [312]

Li, Q., and Racine, J. (2008), "Nonparametric Estimation of Conditional cdf and Quantile Functions With Mixed Categorical and Continuous Data," *Journal of Business & Economic Statistics*, 26, 423–434. [309]

Lv, Y., Zhang, R., Zhao, W., and Liu, J. (2015), "Quantile Regression and Variable Selection of Partial Linear Single-Index Model," *Annals of the Institute of Statistical Mathematics*, 67, 375–409. [309]

Ma, S., and He, X. (2016), "Inference for Single-Index Quantile Regression Models With Profile Optimization," *Annals of Statistics*, 44, 1234–1268. [310]

Oh, H.-S., Lee, T., and Nychka, D. (2011), "Fast Nonparametric Quantile Regression With Arbitrary Smoothing Methods," *Journal of Computational and Graphical Statistics*, 20, 510–526. [309]

Santos, B., and Kneib, T. (2020), "Noncrossing Structured Additive Multiple-Output Bayesian Quantile Regression Models," *Statistics and Computing*, 30, 855–869. [310]

Schnabel, S., and Eilers, P. (2013), "Simultaneous Estimation of Quantile Curves Using Quantile Sheets," *AStA Advances in Statistical Analysis*, 97, 77–87. [310]

Silverman, B. (1986), *Density Estimation*, London: Chapman and Hall. [314]

Takeuchi, I., Le, Q., Sears, T., and Smola, A. (2006), "Nonparametric Quantile Regression," *Journal of Machine Learning Research*, 7, 1231–1264. [309]

Tang, Y., Wang, H., and Liang, H. (2018), "Composite Estimation for Single-Index Models With Responses Subject to Detection Limits," *Scandinavian Journal of Statistics*, 45, 444–464. [310]

Wu, T., Yu, K., and Yu, Y. (2010), "Single-Index Quantile Regression," *Journal of Multivariate Analysis*, 101, 1607–1621. [309,310,311,312,313,316]

Wu, Y., and Liu, Y. (2009), "Stepwise Multiple Quantile Regression Estimation Using Non-Crossing Constraints," *Statistics and Its Interface*, 2, 299–310. [310,312,313]

Yeh, I.-C., and Hsu, T.-K. (2018), "Building Real Estate Valuation Models With Comparative Approach Through Case-Based Reasoning," *Applied Soft Computing*, 65, 260–271. [319]

Yu, K., and Jones, M. (1998), "Local Linear Quantile Regression," *Journal of the American Statistical Association*, 93, 228–237. [309,310]

Yu, Y., and Ruppert, D. (2002), "Penalized Spline Estimation for Partially Linear Single-Index Models," *Journal of the American Statistical Association*, 97, 1042–1054. [309]

Zhu, L., and Xue, L. (2006), "Empirical Likelihood Confidence Regions in a Partially Linear Single-Index Model," *Journal of the Royal Statistical Society*, Series B, 68, 549–570. [309]

Zou, H., and Yuan, M. (2008), "Composite Quantile Regression and the Oracle Model Selection Theory," *Annals of Statistics*, 36, 1108–1126. [310]