Ann Inst Stat Math (2017) 69:145–168
DOI 10.1007/s10463-015-0538-5

Composite change point estimation for bent line quantile regression

Liwen Zhang[1] · Huixia Judy Wang[3] · Zhongyi Zhu[2]

Received: 7 May 2014 / Revised: 9 February 2015 / Published online: 15 August 2015
© The Institute of Statistical Mathematics, Tokyo 2015

Abstract The bent line quantile regression describes the situation where the conditional quantile function of the response is piecewise linear but still continuous in covariates. In some applications, the change points at which the quantile functions are bent tend to be the same across quantile levels or for quantile levels lying in a certain region. To capture such commonality, we propose a composite estimation procedure to estimate model parameters and the common change point by combining information across quantiles. We establish the asymptotic properties of the proposed estimator, and demonstrate the efficiency gain of the composite change point estimator over that obtained at a single quantile level through numerical studies. In addition, three different inference procedures are proposed and compared for hypothesis testing and the construction of confidence intervals. The finite sample performance of the proposed procedures is assessed through a simulation study and the analysis of a real data.

Keywords Change point · Composite quantile regression · Bent line · Bootstrap · Rank score

Zhang's research is partially supported by NNSFC Grant 11171074. Wang's research is partially supported by NSF (National Science Foundation) CAREER Award DMS-1149355. Zhu's research is partially supported by NSFC 11271080.

✉ Zhongyi Zhu
zhuzy@fudan.edu.cn

1 School of Economics, Shanghai University, Shanghai 200444, China

2 Department of Statistics, Fudan University, Shanghai 200433, China

3 Department of Statistics, George Washington University, Washington, DC 20052, USA

1 Introduction

In conventional regression models, the response is often modeled as a single parametric function on the whole domain of the predictors. However, in some applications, this assumption of the stability of the regression coefficients is not satisfied, and instead it is more appropriate to consider a threshold model with a change point at an unknown threshold in one covariate; see Vieth (1989), Pastor and Guallar (1998) and Fiteni (2004). In this paper, we focus on a special case of the threshold model, the bent line regression, also referred to as broken line regression or segmented linear regression model. In bent line regression, the regression function has different linear forms in different regions of the threshold covariate with continuity at the change point. One can refer to Feder (1975), Chappell (1989), Muggeo (2003) and Li et al. (2011) for a comprehensive review and various applications.

Numerous work has been done for the estimation and hypothesis testing in segmented regression models. For example, Robinson (1964) discussed maximum likelihood estimation with and without constraints on the change point assuming a normal distribution. Feder (1975) studied the asymptotic distribution theory in segmented least squares regression. Liu et al. (1997) studied multiple-segment least squares regression, where a modified Schwarz criterion was used to determine the number of segments. Chan and Tsay (1998) studied the limiting properties of least squares estimator in a two-phase threshold autoregressive model, which includes bent line regression as a special case. Muggeo (2003) developed an iterative estimation algorithm based on a simple linearization technique for fitting the piecewise terms in regression models with unknown change points. Liu and Qian (2010) proposed an empirical likelihood method for hypothesis testing and estimation in a two-phase segmented least squares regression model. Kosorok and Song (2007) studied estimation and hypothesis testing for segmented transformation models applied to right censored survival data.

Most existing work for segmented regression focused on modeling the mean function of the response variable. However, in some applications such as the studies of blood pressure and birth weight, the upper or lower quantiles of the response variable are of more scientific interest. Quantile regression, first developed by Koenker and Bassett (1978), provides a natural and flexible way to capture the relationship between the response and covariates at different locations of the response distribution.

There exists limited work for quantile regression with change points. Su and Xiao (2008), Qu (2008) and Oka and Qu (2011) discussed the hypothesis testing and estimation of change points in quantile regression for time series data, where the structural change of the regression function is due to time change instead of covariate threshold effect. Lee et al. (2011) developed a sup-likelihood-ratio-type method and Zhang et al. (2014) developed a sup-score-type method for testing the existence of a covariate threshold effect in regression models including quantile regression. Galvao et al. (2014) developed a uniform test based on the supremum of the Wald process for testing the linearity against threshold effects in quantile regression. Li et al. (2011) proposed a change point estimation approach for bent line quantile regression, where the change point is estimated at each given quantile level separately. The maximal running speed example in Li et al. (2011) suggested that the structural changes can be heterogenous in the sense that the changes in the regression slopes may have different magnitudes at different quantiles, but they appeared to occur around the same location across quantiles. In such situations, it is more informative to consider joint modeling of multiple quantiles.

In this paper, we propose a composite change point estimator for bent line quantile regression, where information from a range of quantiles is combined to estimate the common change point. Compared to the literature, this paper has the following main contributions. First, by accommodating the commonality of change points, the proposed joint modeling of multiple quantiles often leads to more efficient estimation than the method based on a single quantile level (Li et al. 2011). To our knowledge, this is the first time that a composite estimator is considered for bent line quantile regression. Second, by adopting noncrossing constraints, the proposed method ensures that the estimated conditional quantile functions are nondecreasing in the quantile level. Third, besides establishing the asymptotic properties of the proposed change point and quantile coefficient estimators, we also develop and compare three approaches, Wald-type, bootstrap and rank-score-based, for hypothesis testing and confidence interval construction.

The rest of the paper is organized as follows. In Sect. 2, we describe the proposed composite change point estimator obtained by joint modeling of multiple quantiles for bent line regression. In addition, we present the asymptotic properties of the proposed estimator, and discuss three hypothesis testing procedures. The finite sample performance of the proposed method is investigated through a simulation study in Sect. 3 and the analysis of a blood pressure data set in Sect. 4. All the technical proofs are given in the "Appendix".

2 Estimating the common change point

2.1 Model and the proposed estimation method

Let Y be the response variable of interest. In addition, let X be the univariate threshold variable, and $\mathbf{Z}$ be a q-dimensional vector of covariates. Denote $\mathbf{W} = (X, \mathbf{Z}^T)^T$. Throughout the paper, we assume that the threshold variable X has a bounded support $[M_1, M_2]$, where $M_1 < M_2$ are constants. At any given quantile level $\tau \in (0, 1)$, the τth conditional quantile of Y given $\mathbf{W}$ is defined as $Q_Y(\tau|\mathbf{W}) = F^{-1}(\tau|\mathbf{W}) = \inf\{t : F(t|\mathbf{W}) \geq \tau\}$, where $F(\cdot|\mathbf{W})$ is the conditional distribution function of Y given $\mathbf{W}$.

Suppose that the threshold variable X has segmented effects on the quantiles of Y in the interval $\mathcal{T} = [\omega_1, \omega_2]$ with $0 < \omega_1 < \omega_2 < 1$. We consider the following bent line quantile regression model

$$\begin{aligned} Q_Y(\tau; \boldsymbol{\eta}_{\tau,0}, u_0|\mathbf{W}) &= \alpha_{\tau,0} + \beta_{1,\tau,0}(X - u_0)I(X \leq u_0) \\ &\quad + \beta_{2,\tau,0}(X - u_0)I(X > u_0) + \mathbf{Z}^T\boldsymbol{\gamma}_{\tau,0}, \quad \text{for } \tau \in \mathcal{T}, \end{aligned} \tag{1}$$

where $\boldsymbol{\eta}_{\tau,0} = (\alpha_{\tau,0}, \beta_{1,\tau,0}, \beta_{2,\tau,0}, \boldsymbol{\gamma}_{\tau,0}^T)^T$, $u_0 \in (M_1, M_2)$ is the unknown change point, $\alpha_{\tau,0}$ is the baseline intercept, $\beta_{1,\tau,0} \neq \beta_{2,\tau,0}$ are the coefficients of X before and after the change point, and $\boldsymbol{\gamma}_{\tau,0}$ is the effect of $\mathbf{Z}$. In Model (1), the covariate effects are allowed to vary at different quantiles, but the change point is assumed to be the same for all $\tau \in \mathcal{T}$. If $\mathcal{T} = (0, 1)$, the model implies that the change point is a constant across all quantile levels.

Suppose we observe a random sample $\{(Y_i, X_i, \mathbf{Z}_i)\}_{i=1}^n$ of $(Y, X, \mathbf{Z})$. Let $\tau_1 \leq \cdots \leq \tau_K$ be a set of quantile levels in $\mathcal{T}$. Denote $\boldsymbol{\eta}_k = (\alpha_{\tau_k}, \beta_{1,\tau_k}, \beta_{2,\tau_k}, \boldsymbol{\gamma}_{\tau_k}^T)^T$ for $k = 1, \ldots, K$, $\boldsymbol{\eta} = (\boldsymbol{\eta}_1^T, \ldots, \boldsymbol{\eta}_K^T)^T$ and $\boldsymbol{\theta} = (\boldsymbol{\eta}^T, u)^T$. Define

$$S_n(\boldsymbol{\eta}, u) = n^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n} \rho_{\tau_k}\{Y_i - Q_Y(\tau_k; \boldsymbol{\eta}, u|\mathbf{W}_i)\}, \tag{2}$$

where $\rho_\tau(v) = v\{\tau - I(v < 0)\}$ is the quantile loss function, and $I(\cdot)$ is the indicator function. Throughout, let $\boldsymbol{\theta}_0$ denote the true value of θ, we propose to estimate $\boldsymbol{\theta}_0$ by

$$\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\eta}}^T, \hat{u})^T = \underset{\eta \in \mathcal{B}, u \in (M_1, M_2)}{\operatorname{argmin}} S_n(\boldsymbol{\eta}, u), \tag{3}$$

where $\mathcal{B} \subset \mathbb{R}^{K(q+3)}$ is a compact set.

Different from the quantile-specific estimator in Li et al. (2011), our proposed change point estimator is obtained by combining information across quantiles. From now on, we refer to the proposed change point estimator $\hat{u}$ as the composite change point estimator. Composite estimators have been considered in different settings for models without change points; see for instance Zou and Yuan (2008), Kai et al. (2010, 2011), Jiang et al. (2012, 2013). When the change point is indeed constant, the composite change point estimator often leads to higher efficiency than the quantile-specific estimator.

Similar to Chan and Tsay (1998) and Li et al. (2011), we solve the minimization problem in (3) using a two-stage profile procedure. First, define the profile estimator of $\boldsymbol{\eta}$ at a given candidate u as

$$\hat{\boldsymbol{\eta}}(u) = \underset{\eta \in \mathcal{B}}{\operatorname{argmin}}\, S_n(\boldsymbol{\eta}, u). \tag{4}$$

For analysis at multiple quantiles, the estimated quantile curves may cross, yielding a larger quantile estimate at a lower quantile than that at an upper quantile. Quantile crossing issues have been discussed in He (1997), Dette and Volgushev (2008), Chernozhukov et al. (2010), Bondell et al. (2010), among others, for regression models without change points. To avoid quantile crossing, we adopt the noncrossing constraint of Bondell et al. (2010) in the estimation of $\boldsymbol{\eta}(u)$ using the R function at http://www4.stat.ncsu.edu/~hdbondel/Software/NoCross.

In the second stage, the common change point u_0 can be estimated by

$$\hat{u} = \underset{u \in [M_1 + \varepsilon_1, M_2 - \varepsilon_2]}{\operatorname{argmin}} S_n\{\hat{\boldsymbol{\eta}}(u), u\},$$

where ε_1 and ε_2 are two small positive constants. The trimming parameters ε_1 and ε_2 are included to avoid searching u on the boundaries of X, which corresponds to the regression model with no structural changes and thus the change point is not identifiable. The final estimator of $\boldsymbol{\theta}$ is obtained by $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\eta}}(\hat{u})^T, \hat{u})^T$. With the noncrossing constrained estimator of $\boldsymbol{\eta}(u)$, the proposed estimation procedure ensures that the estimated quantile function $Q_Y(\tau; \hat{\boldsymbol{\eta}}(u), u|\mathbf{W}_i)$ is nondecreasing in τ at any candidate u and thus at $\hat{u}$. In addition, the asymptotic properties of the estimator $\hat{\boldsymbol{\theta}}$ are not affected by the noncrossing constraints.

In practice, we need to choose $\mathcal{T}$ and $\tau_k, k = 1, \cdots, K$. The interval $\mathcal{T}$ is often determined given the problem of interest. In our context, we also need to carry out hypothesis testing to assess the constancy of change points across quantiles in $\mathcal{T}$. The Wald-type test suggested in Li et al. (2011) for general linear hypothesis testing can be adapted for this purpose. Once $\mathcal{T}$ is determined, the quantile levels τ_k can be chosen uniformly spaced within $\mathcal{T}$. For linear quantile regression without change points, Koenker (1984) showed that as $K \to \infty$, the optimally weighted composite quantile regression estimator can achieve the same efficiency as the maximum likelihood estimator. However, a larger number of K is associated with higher computational cost as more unknown parameters are involved. Evidence from our empirical studies suggests that letting $K \geq 9$ often gives little additional efficiency gain; see a detailed sensitivity analysis in Sect. 3.

2.2 Asymptotic properties

We first introduce some notations. Denote $\hat{\boldsymbol{\eta}}_k = (\hat{\alpha}_{\tau_k}, \hat{\beta}_{1,\tau_k}, \hat{\beta}_{2,\tau_k}, \hat{\boldsymbol{\gamma}}_{\tau_k}^T)^T$ and $\hat{\boldsymbol{\eta}} = (\hat{\boldsymbol{\eta}}_1^T, \ldots, \hat{\boldsymbol{\eta}}_K^T)^T$. Let $\boldsymbol{\eta}_{k,0} = (\alpha_{\tau_k,0}, \beta_{1,\tau_k,0}, \beta_{2,\tau_k,0}, \boldsymbol{\gamma}_{\tau_k,0}^T)^T$, $\boldsymbol{\eta}_0 = (\boldsymbol{\eta}_{1,0}^T, \ldots, \boldsymbol{\eta}_{K,0}^T)^T$ and $\boldsymbol{\theta}_0 = (\boldsymbol{\eta}_0^T, u_0)^T$ denote the corresponding true parameter vectors. Throughout the paper, we use $\|\mathbf{z}\|$ to denote the Euclidean norm of any vector $\mathbf{z}$. Furthermore, define

$$\boldsymbol{C}_n = n^{-1} \sum_{j=1}^{K} \sum_{k=1}^{K} \sum_{i=1}^{n} \{\min(\tau_j, \tau_k) - \tau_j \tau_k\} E\{\boldsymbol{h}_j(\mathbf{W}_i; \boldsymbol{\theta}_0) \boldsymbol{h}_k^T(\mathbf{W}_i; \boldsymbol{\theta}_0)\}, \quad (5)$$

and

$$\boldsymbol{D}_n = n^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n} \frac{\partial}{\partial \boldsymbol{\theta}} E\left[\psi_{\tau_k}\{Y_i - Q_Y(\tau_k; \boldsymbol{\theta}|\mathbf{W}_i)\} \boldsymbol{h}_k(\mathbf{W}_i; \boldsymbol{\theta})\right]\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0},$$

where $\boldsymbol{h}_k(\mathbf{W}_i; \boldsymbol{\theta}) = (\mathbf{0}_{(k-1)(q+3)}^T, \boldsymbol{m}^T(\mathbf{W}_i; u), \mathbf{0}_{(K-k)(q+3)}^T, -\beta_{1,\tau_k} I(X_i \leq u) - \beta_{2,\tau_k} I(X_i > u))^T$, $\boldsymbol{m}(\mathbf{W}_i; u) = (1, (X_i - u) I(X_i \leq u), (X_i - u) I(X_i > u), \mathbf{Z}_i^T)^T$, $\psi_\tau(v) = \tau - I(v \leq 0)$, and $\mathbf{0}_q$ is a q-dimensional vector of zeros.

To establish the asymptotic distribution of the proposed estimator, we make the following assumptions.

- A1 For each $i = 1, \ldots, n$ and $k = 1, \ldots, K$, $F_i \equiv F(\cdot|\mathbf{W}_i)$ has a continuous density $f_i(\cdot)$ that is uniformly bounded away from 0 and ∞ at the points $F^{-1}(\tau_k|\mathbf{W}_i)$.
- A2 The density function of X_i is continuous with a compact support $[M_1, M_2]$.
- A3 $E(\|\mathbf{Z}_i\|^3)$ is bounded.

A4 There exist two positive definite matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ such that $\lim_{n\to\infty} \boldsymbol{C}_n = \boldsymbol{C}$ and $\lim_{n\to\infty} \boldsymbol{D}_n = \boldsymbol{D}$.

Assumption A1 is standard in quantile regression. Assumptions A2–A3 give some conditions on the covariates. Assumptions A1–A3 are needed to show the consistency of $\hat{\boldsymbol{\theta}}$. Assumption A4 is an additional condition needed to establish the asymptotic normality of $\hat{\boldsymbol{\theta}}$.

The following theorem states the asymptotic normality of the proposed estimator of $\boldsymbol{\theta}_0$.

Theorem 1 *Suppose that Model (1) and Assumptions A1–A4 hold, as $n \to \infty$, we have*

$$n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \boldsymbol{\Sigma}),$$

where $\boldsymbol{\Sigma} = \boldsymbol{D}^{-1}\boldsymbol{C}\boldsymbol{D}^{-1}$.

2.3 Construction of confidence intervals

In this section, we propose three different procedures for constructing confidence intervals of elements of $\boldsymbol{\theta}_0$, including Wald-type, bootstrap and rank-score-based methods. The Wald-type and rank-score-based methods are large-sample inference procedures, while the performance of the former method is more sensitive to the estimation of the error density function in finite samples.

2.3.1 Wald-type method

Based on the asymptotic normality in Theorem 1, an asymptotic confidence interval for any element of $\boldsymbol{\theta}_0$ can be constructed by directly estimating the covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{D}^{-1}\boldsymbol{C}\boldsymbol{D}^{-1}$. The matrix $\mathbf{C}$ can be estimated consistently by

$$\hat{\mathbf{C}}_n = n^{-1} \sum_{j=1}^{K} \sum_{k=1}^{K} \sum_{i=1}^{n} \{\min(\tau_j, \tau_k) - \tau_j \tau_k\} h_j(\mathbf{W}_i; \hat{\boldsymbol{\theta}}) h_k^T(\mathbf{W}_i; \hat{\boldsymbol{\theta}}).$$

The estimation of $\mathbf{D}$ is more complicated, since the matrix involves the unknown error densities $f_i\{Q_Y(\tau_k; \boldsymbol{\theta}_0|\mathbf{W}_i)\}$, $k = 1, \ldots, K$. Adopting the difference quotient idea in Hendricks and Koenker (1992), we propose to estimate $f_i\{Q_Y(\tau_k; \boldsymbol{\theta}_0|\mathbf{W}_i)\}$ by

$$\hat{f}_i\{Q_Y(\tau_k; \boldsymbol{\theta}_0|\mathbf{W}_i)\} = \frac{2\Delta_{n,k}}{Q_Y(\tau_k + \Delta_{n,k}, \hat{\boldsymbol{\theta}}|\mathbf{W}_i) - Q_Y(\tau_k - \Delta_{n,k}, \hat{\boldsymbol{\theta}}|\mathbf{W}_i)}, \tag{6}$$

where $\Delta_{n,k}$ is a bandwidth parameter going to zero as $n \to \infty$, and $Q_Y(\tau_k \pm \Delta_{n,k}; \hat{\boldsymbol{\theta}}|\mathbf{W}_i)$ are the estimated $(\tau_k \pm \Delta_{n,k})$th conditional quantiles of Y_i. In our empirical studies, following the suggestion in Hall and Sheather (1988) we choose $\Delta_{n,k}$ based on Edgeworth expansions of studentized quantiles by

$$\Delta_{n,k} = 1.57n^{-1/3}(1.5\phi^2\{\Phi^{-1}(\tau_k)\}/[2\{\Phi^{-1}(\tau_k)\}^2 + 1])^{2/3},$$

where ϕ and Φ are the density and distribution functions of the standard normal distribution.

Therefore, $\boldsymbol{D}$ can be consistently estimated by

$$\hat{\boldsymbol{D}}_n = \begin{pmatrix} -\hat{\boldsymbol{D}}_{n11} & \hat{\boldsymbol{D}}_{n12} \\ \hat{\boldsymbol{D}}_{n12}^T & -\hat{D}_{n22} \end{pmatrix},$$

where $\hat{\boldsymbol{D}}_{n11} = \text{diag}(\hat{\boldsymbol{D}}_{n11,1}, \ldots, \hat{\boldsymbol{D}}_{n11,K})$ with $\hat{\mathbf{D}}_{n11,k} = n^{-1}\sum_{i=1}^n \hat{f}_i\{Q_Y(\tau_k; \boldsymbol{\theta}_0 |\mathbf{W}_i)\}\boldsymbol{m}(\mathbf{W}_i; \hat{u})\boldsymbol{m}^T(\mathbf{W}_i; \hat{u})$, $\hat{\boldsymbol{D}}_{n12} = (\hat{\boldsymbol{D}}_{n12,1}^T, \ldots, \hat{\boldsymbol{D}}_{n12,K}^T)^T$ with $\hat{\boldsymbol{D}}_{n12,k} = n^{-1} \sum_{i=1}^n \hat{f}_i\{Q_Y(\tau_k; \boldsymbol{\theta}_0|\mathbf{W}_i)\}\{\hat{\beta}_{1,\tau_k} I(X_i \le \hat{u}) + \hat{\beta}_{2,\tau_k} I(X_i > \hat{u})\}\boldsymbol{m}(\mathbf{W}_i; \hat{\boldsymbol{\theta}})$ and $\hat{D}_{n22} = n^{-1}\sum_{k=1}^K\sum_{i=1}^n \hat{f}_i\{Q_Y(\tau_k; \boldsymbol{\theta}_0|\mathbf{W}_i)\}\{\hat{\beta}_{1,\tau_k}^2 I(X_i \le \hat{u}) + \hat{\beta}_{2,\tau_k}^2 I(X_i > \hat{u})\}$.

2.3.2 Bootstrap method

The Wald-type method is easy to implement, but its finite sample performance could be sensitive to the choice of bandwidth $\Delta_{n,k}$; see for instance in Kocherginsky et al. (2005) for comparison of different inference methods for linear quantile regression. Our simulation study in Sect. 3 shows that the directly estimated variance tends to be underestimated in small samples for the scenarios considered, leading to confidence intervals with coverage lower than the nominal level.

We consider a simple paired bootstrap method (Freedman (1981)). The method works by first resampling the paired observations $\{(Y_i, \mathbf{W}_i)\}_{i=1}^n$ with replacement, and then calculating the bootstrap estimator $\{\hat{\boldsymbol{\theta}}_b^*, b = 1, \ldots, B\}$ based on the bootstrap samples, where B is the number of bootstrap repetition. The confidence interval for any element of $\boldsymbol{\theta}_0$ can then be constructed by the sample percentiles of the bootstrap estimates.

2.3.3 Rank score test

The direct variance estimation method involves estimating the unknown conditional density functions, while the bootstrap method is computationally intensive. In linear quantile regression, rank score test was shown to be a stable alternative inference procedure; see for instance Gutenbrunner et al. (1993), Koenker (1994), Chen and Wei (2005), Kocherginsky et al. (2005), Wang and He (2007) and Wang et al. (2009).

We adapt the rank score method to inference for bent line quantile regression. We focus on constructing confidence intervals for the common change point u_0, since this is the parameter of main interest in this paper. Let u_{τ_k} denote the change point associated with the τ_kth quantile, $k = 1, \ldots, K$. We first consider the following hypotheses

$$H_0 : u_{\tau_1} = \cdots = u_{\tau_K} = u_0 \text{ v.s. } H_1 : u_{\tau_k} \ne u_0 \text{ for some } 1 \le k \le K. \tag{7}$$

The null hypothesis states that the change points at all K quantiles are the same as u_0.

Under H_0, for each quantile τ_k, the regression coefficients $\boldsymbol{\eta}_k$ can be estimated by fitting the segmented quantile regression model (1) with the known change point u_0 using existing software. For $k = 1, \ldots, K$, we denote the resulting estimator of $\boldsymbol{\eta}_k$ as $\hat{\boldsymbol{\eta}}_k(u_0)$, and the corresponding estimated residuals as $\hat{e}_{i,\tau_k} = Y_i - Q_Y\{\tau_k; \hat{\boldsymbol{\eta}}_k(u_0), u_0|\mathbf{W}_i\}$, $i = 1, \ldots, n$. Define $\mathbf{M}(u) = (m(\mathbf{W}_1; u), \ldots, m(\mathbf{W}_n; u))^T$ and $\mathbf{P}_k = (p_1\{\tau_k; \hat{\boldsymbol{\eta}}_k(u_0), u_0\}, \ldots, p_n\{\tau_k; \hat{\boldsymbol{\eta}}_k(u_0), u_0\})^T$, where $p_i(\tau_k; \boldsymbol{\eta}_k, u_0) = -\beta_{1,\tau_k} I(X_i \leq u_0) - \beta_{2,\tau_k} I(X_i > u_0)$ is the first derivative of $Q_Y(\tau_k; \boldsymbol{\eta}_k, u|\mathbf{W}_i)$ with respect to u evaluated at $u = u_0$. In addition, define $\mathbf{P}_k^* \equiv (p_1^*\{\tau_k; \hat{\boldsymbol{\eta}}_k(u_0), u_0\}, \ldots, p_n^*\{\tau_k; \hat{\boldsymbol{\eta}}_k(u_0), u_0\})^T = (\boldsymbol{I}_n - \boldsymbol{\Lambda}_k)\boldsymbol{P}_k$ with $\boldsymbol{\Lambda}_k = \mathbf{M}(u_0)\{\mathbf{M}^T(u_0)\boldsymbol{B}_k\,\mathbf{M}(u_0)\}^{-1}\mathbf{M}^T(u_0)\boldsymbol{B}_k$, where $\boldsymbol{B}_k = diag(\hat{f}_1\{Q_Y(\tau_k; \boldsymbol{\eta}_{k,0}, u_0|\mathbf{W}_1)\}, \ldots, \hat{f}_n\{Q_Y(\tau_k; \boldsymbol{\eta}_{k,0}, u_0|\mathbf{W}_n)\})$ with $\hat{f}_i\{Q_Y(\tau_k; \boldsymbol{\eta}_{k,0}, u_0|\mathbf{W}_i)$ defined in (6).

The proposed rank score test statistic is defined as

$$T_n = \boldsymbol{S}_n^T \boldsymbol{V}_n^{-1} \boldsymbol{S}_n,$$

where

$$\boldsymbol{S}_n = (S_{n,1}, \ldots, S_{n,K})^T, \;\; S_{n,k} = n^{-1/2} \sum_{i=1}^{n} p_i^*\{\tau_k; \hat{\boldsymbol{\eta}}_k(u_0), u_0\}\psi_{\tau_k}(\hat{e}_{i,\tau_k}),$$

for $k = 1, \ldots, K$ and $\boldsymbol{V}_n$ is a $K \times K$ matrix with the (k, k')th element equal to

$$n^{-1} \sum_{i=1}^{n} \psi_{\tau_k}(\hat{e}_{i,\tau_k})\psi_{\tau_{k'}}(\hat{e}_{i,\tau_{k'}}) p_i^*\{\tau_k; \hat{\boldsymbol{\eta}}_k(u_0), u_0\} p_i^*\{\tau_{k'}; \hat{\boldsymbol{\eta}}_{k'}(u_0), u_0\}.$$

Before presenting the asymptotic distribution of T_n under H_0, we impose the following additional assumptions.

- A5 For $i = 1, \ldots, n$, F_i has a Lebesgue density that has a bounded first-order derivative.
- A6 The minimum eigenvalue of $\mathbf{V}_n$ is bounded away from zero for sufficiently large n.

The following theorem gives the asymptotic distribution of T_n under H_0.

Theorem 2 *Suppose that Assumptions A1 and A3–A6 hold. Under the null hypothesis H_0 in (7), as $n \to \infty$, we have $T_n \xrightarrow{d} \chi_K^2$.*

Using the asymptotic result in Theorem 2, the rank score test rejects the null hypothesis H_0 when the observed test statistic T_n exceeds $\chi_K^2(1-\alpha)$, the $(1-\alpha)$th quantile of the χ_K^2 distribution, where α is the pre-specified nominal level. By inverting the rank score test, we can construct the $(1-\alpha)$ confidence interval for the common change point u_0 by including all the $\tilde{u}$ for which the null hypothesis $H_0 : u_{\tau_1} = \cdots = u_{\tau_K} = \tilde{u}$ will not be rejected. Below is the proposed procedure for constructing a 95 % confidence interval for u_0.

Step 1. Obtain the composite estimator $\hat{u}$ using the profile estimating method introduced in Sect. 2.1, and calculate the estimated standard error of $\hat{u}$ using the direct variance estimation method described in Sect. 2.3.1, denoted as $\hat{\sigma}_u$.

Step 2. Search for the lower bound of the confidence interval u_l. Define a grid of N points in the left neighborhood of $\hat{u}$: $\mathcal{U}_l = [\hat{u} - 3\hat{\sigma}_u, \hat{u}]$. Test $H_0 : u_{\tau_1} = \cdots = u_{\tau_K} = \tilde{u}$ for any $\tilde{u} \in \mathcal{U}_l$ at the significance level of 0.05. Define $u_l = \min\{\tilde{u} \in \mathcal{U}_l : \tilde{u}$ is not rejected$\}$, that is, the minimum accepted point. One could also define u_l as the maximum rejected point. Theoretically, the two definitions are equivalent. However, in finite samples, the former definition often leads to wider while the later gives narrower confidence intervals. To account for this possible discrepancy, we propose a further adjustment. Due to the convexity of T_n in $\tilde{u}$, ideally points within $[u_l, \hat{u}]$ should all be accepted. However, if u_l is too low, some points in this interval may be rejected. If the number of rejected points within $[u_l, \hat{u}]$ exceeds a pre-specified value $\varrho > 0$, u_l will be replaced by the maximum rejected point and is kept unchanged otherwise.

Step 3. Follow the similar procedure as in Step 2 to search for the upper bound $u_h \in \mathcal{U}_h = [\hat{u}, \hat{u} + 3\hat{\sigma}_u]$. Define $u_h = \max\{\tilde{u} \in \mathcal{U}_h : \tilde{u}$ is not rejected$\}$. If the number of rejected points within $[\hat{u}, u_h]$ exceeds ϱ, we replace u_h by $u_h = \min\{\tilde{u} \in \mathcal{U}_h : \tilde{u}$ is rejected$\}$ and keep it unchanged otherwise.

The parameter ϱ provides a balance between the length and coverage of confidence intervals. A smaller value of ϱ tends to give narrower confidence intervals with lower coverage, and vice versa. In our empirical studies, we let $N = 50$ and choose a conservative value $\varrho = 16$ to ensure the desired coverage probability. This combination was shown to perform reasonably well in Sect. 3.

3 Simulation study

For this simulation study, we consider six different cases. Data for Cases 1–4 are generated from the following model

$$Y_i = 1 + 3(X_i - u_0)I(X_i \le u_0) - 3(X_i - u_0)I(X_i > u_0) + \sigma(X_i)\epsilon_i,$$

where $X_i \sim U(0, 10)$, $\epsilon_i \sim N(0, 1)$ for Cases 1, 2 and 4, and $\epsilon_i \sim t_3$ for Case 3. The function $\sigma(\cdot)$ determines the model heteroscedasticity. We let $\sigma(X_i) = 1$ in Case 1 corresponding to a homoscedastic case, and $\sigma(X_i) = 1 + 0.2X_i$ in Cases 2–4 corresponding to heteroscedastic cases. We let the change point $u_0 = 5$ in Cases 1–3 and $u_0 = 2$ in Case 4 corresponding to scenarios with symmetric and asymmetric change points, respectively.

For Case 5, we consider the following model with an additional predictor

$$Y_i = 1 + 3(X_i - u_0)I(X_i \le u_0) - 3(X_i - u_0)I(X_i > u_0) + 2Z_i + \sigma(X_i, Z_i)\epsilon_i,$$

where $X_i \sim U(0, 10)$, $Z_i \sim U(-5, 5)$, $\sigma(X_i, Z_i) = 1 + 0.2X_i + 0.2Z_i$, $u_0 = 5$ and $\epsilon_i \sim N(0, 1)$.

Table 1 Average bias (bias) and mean squared error (MSE), multiplied by a factor of 100, of different change point estimators

Case	LSE		LAD		CQRNC	
	Bias	MSE	Bias	MSE	Bias	MSE
$n = 200$						
1. Normal-HO	−0.21	0.26	−0.02	0.41	−0.15	0.28
2. Normal-HE	−0.18	1.15	0.21	1.66	−0.09	1.19
3. T-HE	2.67	4.24	0.29	2.21	0.74	1.81
4. Normal-HE-AS	0.70	0.80	0.98	1.38	0.60	0.83
5. Normal-HE-Z	0.85	1.18	1.21	1.30	0.95	0.92
6. Normal-HE-AS-SE	21.36	75.95	30.92	109.81	15.66	49.50
$n = 500$						
1. Normal-HO	0.08	0.10	0.03	0.15	0.14	0.11
2. Normal-HE	0.18	0.43	0.04	0.63	0.30	0.44
3. T-HE	1.33	1.38	0.70	0.78	0.81	0.67
4. Normal-HE-AS	0.11	0.33	0.12	0.43	0.22	0.31
5. Normal-HE-Z	−0.01	0.43	−0.10	0.46	0.04	0.33
6. Normal-HE-AS-SE	1.81	6.56	1.83	4.89	1.38	3.39

1, Normal-HO: Case 1 with homoscedastic normal errors; 2, Normal-HE: Case 2 with heteroscedastic normal errors; 3, T-HE: Case 3 with heteroscedastic t_3 errors; 4, Normal-HE-AS: Case 4 with heteroscedastic normal error and a change point asymmetric from the center; 5, Normal-HE-Z: Case 5 with heteroscedastic normal errors and an additional predictor Z; 6, Normal-HE-AS-SE: Case 6 with heteroscedastic normal errors, a change point asymmetric from the center and smaller threshold effect

For Case 6, data are generated from the following model

$$Y_i = 1 + (X_i - u_0)I(X_i \leq u_0) - (X_i - u_0)I(X_i > u_0) + \sigma(X_i)\epsilon_i,$$

where $X_i \sim U(0, 10)$, $\sigma(X_i) = 1 + 0.2X_i$, $u_0 = 2$ and $\epsilon_i \sim N(0, 1)$. The threshold effect (i.e., change in the slope before and after the change point) in Case 6 is smaller than those in other cases.

We compare three change point estimators: the proposed composite quantile estimator with noncrossing constraints (CQRNC), the least squares estimator (LSE) proposed in Chan and Tsay (1998), and the least absolute deviation (LAD) estimator of change point proposed in Li et al. (2011). For the composite estimator, we set $K = 9$ and let the equally spaced quantile levels be $\tau_k = k/10,\ k = 1, \ldots, 9$. To avoid searching change points on the boundary, we let the trimming parameters $\varepsilon_1 = \varepsilon_2 = 1$ for all three methods. We consider two sample sizes $n = 200$ and 500. The simulation is repeated 500 times for each scenario.

Table 1 summarizes the average bias and mean squared error of three change point estimators. All estimators have ignorable biases. For Cases 1, 2 and 4 with normal errors, the composite estimator has efficiency comparable to that of LSE. For Cases 5 and 6 with heteroscedastic normal errors, the CQRNC estimator shows higher efficiency than the LSE. The advantage of composite estimator over LSE is more pronounced in Case 3 when errors follow a heavy-tailed distribution. Case 6 represents a more challenging case with weaker threshold effects and the change point located asymmetric from the center, consequently the change point estimation has larger bias and mean squared error than in other cases. Across all different scenarios considered, the CQRNC estimator exhibits higher efficiency than the LAD estimator, confirming the efficiency gain by pooling information across multiple quantiles.

To assess the performance of the proposed inference methods, we report in Table 2 the coverage probabilities and average lengths of 95 % confidence intervals for the change point, constructed by the Wald, bootstrap (Boot) and inversion of rank score (Score) methods. The bootstrap repetition B is set to be 500. For the comparison of computational efficiency, we also present the average CPU time (in seconds) used to construct one confidence interval using the same computer. Generally speaking, the coverage probabilities of the Wald-type confidence intervals are lower than the nominal level for both $n = 200$ and $n = 500$, and the coverage probabilities increase to close 95 % when $n \geq 1000$ (additional simulation not shown due to space limit). Both bootstrap and rank score methods give confidence intervals with coverage probabilities close to 95 %, but the latter leads to slightly wider intervals with significantly much less computing time. In finite samples, we observe that the bootstrap method gives wider confidence intervals than Wald and rank score methods in Case 6, where the change point is asymmetric and threshold effect is small. When the sample size is larger ($n \geq 1000$), confidence intervals from three methods have similar widths (results not shown for space reasons).

The proposed composite estimator requires choosing K, the number of quantile levels. To examine the sensitivity of the composite change point estimator to K, we report in Table 3 the mean squared errors of the composite change point estimator obtained by combining information from quantiles $\tau_k = k/(K+1), k = 1, \ldots, K$, where $K = 1, 3, 6, 9, 19$ and 29. The estimator with $K = 1$ is based on $\tau = 0.5$ and thus is equivalent to the LAD estimator in Table 1. Generally speaking, combining information across quantiles as few as three gives more efficient estimation than using a single quantile. The efficiency of the composite estimator tends to increase with K initially, but it becomes stable when $K \geq 9$. This suggests that in finite samples it suffices to consider K around 10 to achieve a good balance between the numerical and computational efficiency.

4 Blood pressure and body mass index analysis

It is well known that high blood pressure (BP) is a serious chronic condition that may lead to coronary heart disease, kidney failure, stroke and many other health problems. One interest in epidemiological research is to understand the association of BP with body mass index (BMI). Even though linear regression has been widely used, many medical studies targeted on different populations found that threshold effect models provide better fits to BP and BMI data. For instance, Bunker et al. (1995) found there exists a threshold at 21.5 kg/m^2 in men and 24.0 kg/m^2 in women for BMI in very lean rural African populations. Kaufman et al. (1997) showed there is a threshold effect at 21 kg/m^2 in women but not for men among the low-BMI populations in Africa and the

Table 2 Coverage probabilities (in percentage) and average lengths of 95 % confidence intervals for change point using the Wald, bootstrap and inversion of rank score methods

Case	Coverage probability			100 × length			CPU time (s)		
	Wald	Boot	Score	Wald	Boot	Score	Wald	Boot	Score
$n = 200$									
1. Normal-HO	92.0	93.6	96.6	19.2	20.3	26.7	2.2	359.9	37.8
2. Normal-HE	92.0	93.4	95.2	38.2	42.3	53.2	2.4	395.8	37.8
3. T-HE	92.4	94.6	96.8	47.9	53.9	66.6	2.4	403.5	37.7
4. Normal-HE-AS	92.0	94.4	98.6	32.6	40.7	48.6	2.7	457.9	26.5
5. Normal-HE-Z	91.2	94.2	95.0	33.5	38.7	47.2	3.3	607.4	33.5
6. Normal-HE-AS-SE	82.6	95.0	92.6	101.3	319.7	151.4	2.7	465.3	26.0
$n = 500$									
1. Normal-HO	93.4	94.2	93.2	12.1	12.3	16.5	4.8	824.2	61.5
2. Normal-HE	93.0	94.0	93.8	24.2	25.1	32.9	5.2	902.2	62.2
3. T-HE	93.2	95.0	94.6	29.8	31.3	40.8	5.3	917.7	61.9
4. Normal-HE-AS	93.2	95.2	96.2	20.9	23.0	29.1	6.0	1059.1	51.7
5. Normal-HE-Z	93.6	95.4	94.2	21.3	22.9	29.3	7.6	1413.3	67.1
6. Normal-HE-AS-SE	91.4	95.2	96.0	63.0	107.1	88.9	6.3	1127.5	51.5

The CPU time summarizes the average time (in seconds) used in computing the confidence interval for one data set

Table 3 Mean squared error (multiplied by a factor of 100) of the composite change point estimator with different K

Case	$K=1$	$K=3$	$K=6$	$K=9$	$K=19$	$K=29$
$n=200$						
1. Normal-HO	0.41	0.30	0.28	0.28	0.27	0.28
2. Normal-HE	1.66	1.28	1.22	1.19	1.19	1.19
3. T-HE	2.21	1.87	1.84	1.81	1.85	1.85
4. Normal-HE-AS	1.38	0.91	0.87	0.83	0.83	0.87
5. Normal-HE-Z	1.30	1.01	0.93	0.92	0.92	0.92
6. Normal-HE-AS-SE	109.81	52.60	48.27	49.50	48.83	48.53
$n=500$						
1. Normal-HO	0.15	0.12	0.11	0.11	0.10	0.10
2. Normal-HE	0.63	0.49	0.45	0.44	0.44	0.43
3. T-HE	0.78	0.71	0.70	0.67	0.68	0.68
4. Normal-HE-AS	0.43	0.34	0.33	0.31	0.31	0.33
5. Normal-HE-Z	0.46	0.35	0.34	0.33	0.32	0.32
6. Normal-HE-AS-SE	4.89	3.65	3.40	3.39	3.34	3.34

Caribbean. In a study of lean rural and semi-urban in West Africa, Kerry et al. (2005) showed that BMI has a nonlinear effect on diastolic BP with a significant "knot" point at BMI equal to 18 kg/m^2 for younger women. All these studies have focused on mean regression. In this section, we analyze a BP and BMI data from the National Health and Nutrition Examination Survey (NHAENES) by quantile regression. The NHAENES was designed to study the health and nutritional status of the adults and children in the United States. We focus on a subset of NHAENES collected from 2005 to 2006, including 789 non-hispanic black males. Different from the analysis in the literature, we aim to examine the impacts of BMI at different locations of the systolic BP distribution using the proposed method for bent line quantile regression.

We carry out a preliminary analysis to identify an appropriate model for lower/central and upper quantiles separately. We first focus on twelve lower and central quantiles $\tau = 0.1, 0.2, \ldots, 0.8, 0.81, \ldots, 0.84$. To examine the existence of a threshold effect of BMI on BP, we employ the sup-likelihood-ratio test procedures proposed in Lee et al. (2011) at each of the twelve quantile levels. In the presence of a BMI threshold effect, at a given quantile level τ, the linear bent line quantile regression model becomes

$$Q_Y(\tau|X_i, Z_i) = \begin{cases} a_1(\tau) + X_i b_1(\tau) + Z_i c(\tau), & X_i \le u(\tau), \\ a_2(\tau) + X_i b_2(\tau) + Z_i c(\tau), & X_i > u(\tau), \end{cases}$$

where X_i and Z_i represent BMI and the age, respectively, $(a_1(\tau), a_2(\tau), b_1(\tau), b_2(\tau), c(\tau))^T$ are the regression coefficients, $u(\tau)$ is the unknown change point at the τth quantile, and $\tau = 0.1, 0.2, \ldots, 0.8, 0.81, \ldots, 0.84$. The lower and upper bounds of the searching interval of the change point are taken to be 10th and 90th percentiles

Table 4 The p values of the sup-likelihood-ratio test for the existence of BMI threshold effect at different quantiles of the BP distribution, and the quantile-specific change point estimates

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.81	0.82	0.83	0.84
LRT	0.01	0.00	0.00	0.02	0.04	0.01	0.00	0.01	0.01	0.01	0.01	0.02
$\hat{u}(\tau)$	19.25	19.71	20.63	20.55	20.51	20.53	20.82	21.74	21.41	21.39	20.82	21.41
τ	0.85	0.86	0.87	0.88	0.89	0.90						
LRT	0.02	0.01	0.00	0.00	0.00	0.00						
$\hat{u}(\tau)$	26.05	26.05	26.25	27.03	26.51	27.11						

of BMI, respectively. We impose a condition $a_2(\tau) = a_1(\tau) + \{b_1(\tau) - b_2(\tau)\}u(\tau)$ for ensuring the continuity of quantile regression model at the point $u(\tau)$. The resulting p values (see Table 4) suggest that BMI has a significant threshold effect in all twelve quantile regression models. Then, following the estimation method in Li et al. (2011), we obtain the quantile-specific change point estimates. Results suggest that the change points at these quantiles are close to each other. Application of the Wald-type test in Li et al. (2011) for testing the constancy of change point across $\tau = 0.1, 0.2, \ldots, 0.8, 0.81, \ldots, 0.84$ yields a p value of 0.447, indicating that change points across these quantiles are constant.

Next, we examine the high quantiles by focusing on six quantiles $\tau = 0.85, 0.86, \ldots, 0.9$. The testing results suggest that there exist significant threshold effects at these quantiles, and the quantile-specific change point estimates tend to be close to each other; see Table 4. In addition, the Wald-type test gives a p value of 0.996.

Therefore, we consider two separate bent line quantile regression models, one assumes a common change point for lower quantiles (LQ) $\tau \in [0.1, 0.84]$, and the other assumes a common change point for high quantiles (HQ) $\tau \in [0.85, 0.9]$.

Let u_1 and u_2 be the common change points under LQ and HQ, respectively. Since we need to separate the quantiles into two parts, we propose the following modified algorithm to estimate the two common change points jointly with noncrossing constraints. We define the joint objective function as

$$S_n(\boldsymbol{\eta}, u_1, u_2) = n^{-1} \sum_{i=1}^{n} \left\{ \sum_{k=1}^{K_1} \rho_{\tau_k}\{Y_i - Q_Y(\tau_k; \boldsymbol{\eta}, u_1|\mathbf{W}_i)\} + \sum_{k=K_1+1}^{K_2} \rho_{\tau_k}\{Y_i - Q_Y(\tau_k; \boldsymbol{\eta}, u_2|\mathbf{W}_i)\} \right\},$$

where $K_1 = 12$ and $K_2 = 18$, $\tau_1 = 0.1$, $\tau_2 = 0.2, \ldots, \tau_8 = 0.8$, $\tau_9 = 0.81$, $\tau_{10} = 0.82, \ldots, \tau_{18} = 0.9$. Then, the regression coefficients and the common change points can be estimated by

Table 5 Parameter estimates and confidence interval estimates from different methods

	LSE	Single		Composite	
		LAD	0.9	LQ	HQ
$\hat{a}_1$	56.18	60.62	81.33	62.39	81.51
$\hat{a}_2$	100.77	100.95	117.78	100.95	112.93
$\hat{b}_1$	2.56	2.35	1.50	2.25	1.50
Wald	[1.29, 3.82]	[1.42, 3.29]	[1.12,1.88]	[1.46, 3.04]	[1.20, 1.79]
Boot	[0.95, 3.50]	[0.90, 3.48]	[1.20, 3.03]	[0.87, 3.22]	[1.20, 2.96]
$\hat{b}_2$	0.38	0.39	0.16	0.39	0.30
Wald	[0.19, 0.56]	[0.13, 0.65]	[−0.31, 0.62]	[0.14, 0.64]	[−0.13, 0.72]
Boot	[0, 0.53]	[−0.07, 0.62]	[−0.35, 0.69]	[−0.09, 0.57]	[−0.22, 0.72]
$\hat{c}$	0.38	0.33	0.67	0.33	0.67
Wald	[−1.55, 2.31]	[−1.61, 2.27]	[−3.51, 4.85]	[0.26, 0.40]	[0.51, 0.83]
Boot	[0.31, 0.45]	[0.25, 0.41]	[0.51, 0.86]	[0.24, 0.39]	[0.51, 0.84]
$\hat{u}$	20.43	20.51	27.11	20.67	26.18
Wald	[18.50, 22.36]	[18.57, 22.45]	[22.93, 31.29]	[19.02, 22.31]	[24.14, 28.22]
Boot	[19.58, 29.77]	[19.04, 32.15]	[19.88, 29.57]	[19.70, 28.75]	[19.82, 29.14]
Rank				[19.72, 21.34]	[25.33, 27.57]

$$\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\eta}}^T, \hat{u}_1, \hat{u}_2)^T = \underset{\boldsymbol{\eta}\in\mathcal{B}, u_j\in(M_1,M_2); j=1,2}{\text{argmin}} S_n(\boldsymbol{\eta}, u_1, u_2),$$

where the minimization is solved using the similar two-stage method as described in Sect. 2.1 with noncrossing constraints.

We summarize the estimations of the regression coefficients and the common change points using the proposed composite method in Table 5, where the estimated regression coefficients $\hat{a}_i$, $\hat{b}_i$, $i = 1, 2$ and $\hat{c}$ correspond to $\tau = 0.5$ for LQ and $\tau = 0.9$ for HQ. For comparison, we also include the estimations from the least squares method and the single quantile method of Li et al. (2011) at $\tau = 0.5$ and 0.9. We also report the 95 % confidence intervals for regression coefficients obtained by the Wald-type and bootstrap methods, and for the common change points using the Wald-type, bootstrap and rank-score-based methods. In addition, we plot the estimated conditional quantile functions of BP against BMI at quantile levels 0.2, 0.5, 0.8 and 0.9 in Fig. 1. It is clear that with the proposed algorithm, the estimated quantile curves do not cross in the covariate region under study.

At all quantiles considered, we observe that the change point estimations under LQ and HQ are asymmetric and the threshold effect is small. Similar to Case 6 in the simulation study, the bootstrap method gives wider confidence intervals than the other two methods. The systolic blood pressure tends to increase with BMI but the relationship becomes weaker for those with BMIs exceeding certain thresholds. Table 5 shows that the change of pattern occurs at BMI 20.67 for lower quantiles $\tau \leq 0.84$.

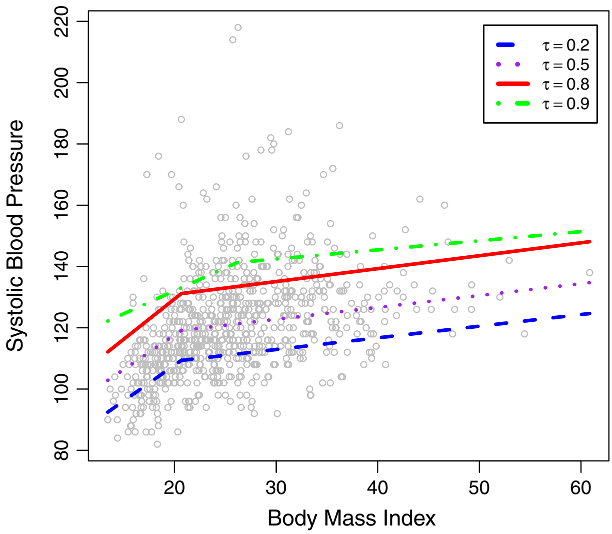

Fig. 1 Estimated quantile regression functions with noncrossing constraints for BP versus BMI for male subjects at the mean age

However, for those at the right tail of the blood pressure distribution, the impact of BMI is not weakened until BMI exceeds 26.18.

Not surprisingly, age shows a positive effect on the blood pressure at both the mean and all the quantiles considered. Comparing to the LSE method and the single quantile analysis approach in Li et al. (2011), the proposed composite method gives shorter confidence intervals for the change points, indicating higher accuracy as observed in the simulation study.

Appendix

Lemma 1 *Suppose Assumptions A1–A3 hold, then* $\hat{\boldsymbol{\theta}}$ *is a consistent estimator of* $\boldsymbol{\theta}_0$.

Proof At a fixed point u, we need to minimize the following objective function

$$n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n}\rho_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\},$$

which is equivalent to minimize

$$n^{-1}\sum_{i=1}^{n}\rho_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\eta},u|\mathbf{W}_i)\},$$

for any $1 \leq k \leq K$. The rest of the proof follows the similar arguments as that of Lemma 1 in Li et al. (2011) and thus is omitted. □

Lemma 2 *Suppose Assumptions A1–A3 hold, we have*

$$\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\le C_2 n^{-1/2}} \Bigg\| n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n}[\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}h_k(\mathbf{W}_i;\boldsymbol{\theta}) - \psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\}h_k(\mathbf{W}_i;\boldsymbol{\theta}_0)] - n^{-1/2}E\Bigg[\sum_{k=1}^{K}\sum_{i=1}^{n}\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\} \times h_k(\mathbf{W}_i;\boldsymbol{\theta})\Bigg]\Bigg\| = o_p(1), \tag{8}$$

where C_2 is some positive constant.

Proof Let

$$u_i(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) = \sum_{k=1}^{K}\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}h_k(\mathbf{W}_i;\boldsymbol{\theta}) - \sum_{k=1}^{K}\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\}h_k(\mathbf{W}_i;\boldsymbol{\theta}_0),$$

where $\mathbf{V}_i$ includes all the random variables Y_i and $\mathbf{W}_i$. Therefore, it can be rewritten as the following

$$u_i(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) = u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) + u_{i2}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) + u_{i3}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) + u_{i4}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0),$$

where $u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) = u_i(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)I\{X_i \le \min(u,u_0)\}$, $u_{i2}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) = u_i(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)I(u_0 < X_i \le u)$, $u_{i3}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) = u_i(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)I(u < X_i \le u_0)$ and $u_{i4}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0) = u_i(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)\ I\{X_i > \max(u,u_0)\}$.

To obtain (8), it is sufficient to show

$$\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|\le C_2 n^{-1/2}} \left\| n^{-1/2}\sum_{i=1}^{n}\{B_i - E(B_i)\}\right\| = o_p(1),$$

where B_i represents $u_{ij}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)$ for $j = 1, 2, 3, 4$. These results follow from Lemma 4.6 in He and Shao (1996), we only show the proof for $B_i = u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)$ for instance. To verify this, we need to check the conditions (B1), (B3) and (B5$'$) of He and Shao (1996).

For (B1), the measurability is easy to show.

For (B3), take $r = 1$, for any $\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\| \le C_2 n^{-1/2}$, we have

$$\|u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)\| = \Bigg\|\Bigg[\sum_{k=1}^{K}\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}h_k(\mathbf{W}_i;\boldsymbol{\theta})$$

$$
\begin{aligned}
&-\sum_{k=1}^{K} \psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\} h_k(\mathbf{W}_i;\boldsymbol{\theta}_0)\Big] I\{X_i \le \min(u,u_0)\}\Big\| \\
&\le \Big\| \sum_{k=1}^{K} \psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}\{h_k(\mathbf{W}_i;\boldsymbol{\theta}) - h_k(\mathbf{W}_i;\boldsymbol{\theta}_0)\} I\{X_i \le \min(u,u_0)\}\Big\| \\
&\quad + \Big\| \sum_{k=1}^{K} [\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\} - \psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\}] h_k(\mathbf{W}_i;\boldsymbol{\theta}_0) \\
&\quad \times I\{X_i \le \min(u,u_0)\}\Big\| \\
&= \|I_{1i}\| + \|I_{2i}\|. \qquad (9)
\end{aligned}
$$

For I_{1i}, it is easy to obtain

$$
E(\|I_{1i}\|^2|\mathbf{W}_i) = n^{-1} O_p(1). \qquad (10)
$$

For I_{2i}, we have

$$
\begin{aligned}
\|I_{2i}\| &= \Big\| \sum_{i=1}^{K} [I\{Y_i \le Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\} - I\{Y_i \le Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}] h_k(\mathbf{W}_i;\boldsymbol{\theta}_0) \\
&\qquad I\{X_i \le \min(u,u_0)\}\Big\| \\
&\le L_1 \|\mathbf{U}_i\| \sum_{k=1}^{K} I\{Q_1(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0) \le Y_i \le Q_2(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0)\} I\{X_i \le \min(u,u_0)\},
\end{aligned}
$$

where L_1 is some constant, $\mathbf{U}_i = (1, X_i, \mathbf{Z}_i^T)^T$, $Q_1(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0) = \min\{Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i), Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}$ and $Q_2(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0) = \max\{Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i),\ Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}$. Thus

$$
\begin{aligned}
&E(\|I_{2i}\|^2|\mathbf{W}_i) \\
&\le L_1^2 \|\mathbf{U}_i\|^2 I\{X_i \le \min(u,u_0)\} E\Big[\sum_{k=1}^{K} \sum_{k'=1}^{K} I\{Q_1(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0) \le Y_i \le Q_2(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0)\} \\
&\quad \times I\{Q_1(\tau_{k'};\boldsymbol{\theta},\boldsymbol{\theta}_0) \le Y_i \le Q_2(\tau_{k'};\boldsymbol{\theta},\boldsymbol{\theta}_0)\} |\mathbf{W}_i \Big].
\end{aligned}
$$

Without loss of generality, we assume $Q_1(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0) < Q_2(\tau_{k'};\boldsymbol{\theta},\boldsymbol{\theta}_0)$ and $Q_1(\tau_{k'};\boldsymbol{\theta},\boldsymbol{\theta}_0) < Q_2(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0)$. Denote $Q_1(\tau_k,\tau_{k'}) = \min\{Q_1(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0), Q_1(\tau_{k'};\boldsymbol{\theta},\boldsymbol{\theta}_0)\}$ and $Q_2(\tau_k,\tau_{k'}) = \max\{Q_2(\tau_k;\boldsymbol{\theta},\boldsymbol{\theta}_0), Q_2(\tau_{k'};\boldsymbol{\theta},\boldsymbol{\theta}_0)\}$. Hence

$$
\begin{aligned}
&E(\|I_{2i}\|^2|\mathbf{W}_i)\\
&\quad \leq L_1^2\|\mathbf{U}_i\|^2\sum_{k=1}^{K}\sum_{k'=1}^{K} f_i(\xi_{k,k'})\{Q_2(\tau_k,\tau_{k'})-Q_1(\tau_k,\tau_{k'})\}I\{X_i\leq \min(u,u_0)\}\\
&\quad \leq L_2 n^{-1/2}\|\mathbf{U}_i\|^3\sum_{k=1}^{K}\sum_{k'=1}^{K} f_i(\xi_{k,k'}),
\end{aligned}
\tag{11}
$$

where the first inequality follows from the mean value theorem with $\xi_{k,k'}$ between $Q_1(\tau_k;\tau_{k'})$ and $Q_2(\tau_k;\tau_{k'})$, the second inequality follows from

$$
\begin{aligned}
&|\{Q_2(\tau_k;\tau_{k'})-Q_1(\tau_k;\tau_{k'})\}I\{X_i\leq \min(u,u_0)\}|\\
&\quad \leq |\{Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)-Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\}I\{X_i\leq \min(u,u_0)\}|\\
&\qquad +|\{Q_Y(\tau_{k'};\boldsymbol{\theta}|\mathbf{W}_i)-Q_Y(\tau_{k'};\boldsymbol{\theta}_0|\mathbf{W}_i)\}I\{X_i\leq \min(u,u_0)\}|\\
&\quad \leq |(\alpha_{\tau_k}-\alpha_{\tau_k,0})+(\boldsymbol{\beta}_{1,\tau_k}-\boldsymbol{\beta}_{1,\tau_k,0})X_i-(\boldsymbol{\beta}_{1,\tau_k}u-\boldsymbol{\beta}_{1,\tau_k,0}u_0)+\mathbf{Z}_i^T(\boldsymbol{\gamma}_{\tau_k}-\boldsymbol{\gamma}_{\tau_k,0})|\\
&\qquad +|(\alpha_{\tau_{k'}}-\alpha_{\tau_{k'},0})+(\boldsymbol{\beta}_{1,\tau_{k'}}-\boldsymbol{\beta}_{1,\tau_{k'},0})X_i-(\boldsymbol{\beta}_{1,\tau_{k'}}u-\boldsymbol{\beta}_{1,\tau_{k'},0}u_0)+\mathbf{Z}_i^T(\boldsymbol{\gamma}_{\tau_{k'}}\\
&\qquad\quad -\boldsymbol{\gamma}_{\tau_{k'},0})|\\
&\quad \leq L_3 n^{-1/2}\|\mathbf{U}_i\|,
\end{aligned}
$$

where L_2 and L_3 are some positive constants satisfying $L_2=L_1^2L_3^2$. By Assumptions A3 and A4, combining (9), (10) and (11), for large n we have

$$
E\{\|u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)\|^2|\mathbf{W}_i)\}\leq Ln^{-1/2}\|\mathbf{U}_i\|^3\sum_{k=1}^{K}\sum_{k'=1}^{K} f_i(\xi_{k,k'}).
$$

It is obvious to obtain (B3) by taking $a_i=\sqrt{L\|\mathbf{U}_i\|^3\sum_{k=1}^{K}\sum_{k'=1}^{K} f_i(\xi_{k,k'})}$.

For (B5$'$), let $A_n=L\sum_{i=1}^{n}\|\mathbf{U}_i\|^3\sum_{k=1}^{K}\sum_{k'=1}^{K} f_i(\xi_{k,k'})$. From Assumptions A2 and A3, we have $E(A_n)=O(n)$. For any positive constant $C_3>0$, taking the decreasing sequence of positive number d_n satisfying $n^{-1/2}(\log n)^4=o(d_n)$ and $d_n=o(1)$, we can show

$$
\begin{aligned}
&P\left(\|\max_{1\leq i\leq n} u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)\|\geq C_3A_n^{1/2}d_n^{1/2}(\log n)^{-2}\right)\\
&\quad \leq \sum_{i=1}^{n} P(\|u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)\|\geq C_3A_n^{1/2}d_n^{1/2}(\log n)^{-2})\\
&\quad \leq \sum_{i=1}^{n}\frac{E\|u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)\|^2}{C_3^2A_nd_n(\log n)^{-4}}\\
&\quad \leq \frac{n^{-1/2}(\log n)^4}{C_3^2d_n}\\
&\quad = o(1).
\end{aligned}
$$

Hence,

$$\max_{1\le i\le n} \|u_{i1}(\mathbf{V}_i;\boldsymbol{\theta},\boldsymbol{\theta}_0)\| = O_p(A_n^{1/2}d_n^{1/2}(\log n)^{-2}),$$

thus (B5') is satisfied. This completes the proof of Lemma 2. □

Proof of Theorem 1 By Lemmas 1 and 2, we obtain

$$\begin{aligned} & n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n}[\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\hat{\boldsymbol{\theta}}|\mathbf{W}_i)\}\boldsymbol{h}_k(\mathbf{W}_i;\hat{\boldsymbol{\theta}}) - \psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\} \\ & \quad \times h_k(\mathbf{W}_i;\boldsymbol{\theta}_0)] - n^{-1/2}\left[E\sum_{k=1}^{K}\sum_{i=1}^{n}\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}\boldsymbol{h}_k(\mathbf{W}_i;\boldsymbol{\theta})\right]\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \\ & \quad = o_p(1). \end{aligned} \tag{12}$$

Applying the Taylor expansion, we get

$$\begin{aligned} & \left[E\sum_{k=1}^{K}\sum_{i=1}^{n}\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}\boldsymbol{h}_k(\mathbf{W}_i;\boldsymbol{\theta})\right]\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = n\boldsymbol{D}_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \\ & \quad + O_p(n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^2), \end{aligned} \tag{13}$$

where

$$\begin{aligned} \boldsymbol{D}_n &= n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n}\frac{\partial E\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}\boldsymbol{h}_k(\mathbf{W}_i;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\ &= n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n}\frac{\partial([\tau_k - F_i\{Q_Y(\tau_k;\boldsymbol{\theta}|\mathbf{W}_i)\}]\boldsymbol{h}_k(\mathbf{W}_i;\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}}\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\ &= n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n}\Bigg([-f_i\{Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\}\boldsymbol{h}_k(\mathbf{W}_i;\boldsymbol{\theta}_0)\boldsymbol{h}_k^T(\mathbf{W}_i;\boldsymbol{\theta}_0)] \\ &\quad + [\tau_k - F_i\{Q(\tau_k)\}]\frac{\partial\boldsymbol{h}_k(\mathbf{W}_i;\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\Bigg) \\ &= n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n}\left[-f_i\{Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\}\boldsymbol{h}_k(\mathbf{W}_i;\boldsymbol{\theta}_0)\boldsymbol{h}_k^T(\mathbf{W}_i;\boldsymbol{\theta}_0)\right]. \end{aligned}$$

In addition, by the subgradient condition of quantile regression (pages 34–38 in Koenker 2005), we have

$$n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n}\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\hat{\boldsymbol{\theta}}|\mathbf{W}_i)\}\boldsymbol{h}_k(\mathbf{W}_i;\hat{\boldsymbol{\theta}}) = o_p(1). \tag{14}$$

Combining (12), (13) and (14), we have

$$-n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n}[\psi_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\}\boldsymbol{h}_k(\mathbf{W}_i;\boldsymbol{\theta}_0)] = n^{1/2}\mathbf{D}_n(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$$
$$+ O_p(n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^2) + o_p(1).$$

This results in $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 = O_p(n^{-1/2})$. Therefore,

$$n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = -\boldsymbol{D}_n^{-1}n^{-1/2}\sum_{k=1}^{K}\sum_{i=1}^{n}\boldsymbol{\psi}_{\tau_k}\{Y_i - Q_Y(\tau_k;\boldsymbol{\theta}_0|\mathbf{W}_i)\}h_k(\mathbf{W}_i;\boldsymbol{\theta}_0) + o_p(1).$$

Following central limit theorem, by Assumption A4 and some simple calculation, we can obtain that $n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ is asymptotically normal with mean zero and variance $\boldsymbol{D}^{-1}\boldsymbol{C}\boldsymbol{D}^{-1}$. This completes the proof of Theorem 1. □

Proof of Theorem 2 Denote

$$T_n^* = \boldsymbol{S}_n^{*T}(V_n^*)^{-1}\boldsymbol{S}_n^*,$$

where $\boldsymbol{S}_n^* = \{S_{n,1}^*, \ldots, S_{n,K}^*\}^T$, $S_{n,k}^* = n^{-1/2}\sum_{i=1}^n p_i^*(\tau_k;\boldsymbol{\eta}_{k,0}, u_0|\mathbf{W}_i)\psi_{\tau_k}(e_{i,\tau_k})$, $e_{i,\tau_k} = Y_i - Q_Y(\tau_k;\boldsymbol{\eta}_{k,0}, u_0|\mathbf{W}_i)$, $\boldsymbol{V}_n^*$ is a $K \times K$ matrix with the $(k, k^{'})$th element

$$\begin{aligned}
&\mathrm{Cov}(S_{n,k}^*, S_{n,k'}^*)\\
&= \mathrm{Cov}\left\{n^{-1/2}\sum_{i=1}^{n}p_i^*(\tau_k;\boldsymbol{\eta}_{k,0},u_0)\psi_{\tau_k}(e_{i,\tau_k}),\ n^{-1/2}\sum_{i=1}^{n}p_i^*(\tau_{k'};\boldsymbol{\eta}_{k',0},u_0)\psi_{\tau_{k'}}(e_{i,\tau_{k'}})\right\}\\
&= n^{-1}\sum_{i=1}^{n}\mathrm{Cov}\{p_i^*(\tau_k;\boldsymbol{\eta}_{k,0},u_0)\psi_{\tau_k}(e_{i,\tau_k}),\ p_i^*(\tau_{k'};\boldsymbol{\eta}_{k',0},u_0)\psi_{\tau_{k'}}(e_{i,\tau_{k'}})\}\\
&= n^{-1}\sum_{i=1}^{n}\{\min(\tau_k,\tau_{k'}) - \tau_k\tau_{k'}\}p_i^*(\tau_k;\boldsymbol{\eta}_{k,0},u_0)p_i^*(\tau_{k'};\boldsymbol{\eta}_{k',0},u_0).
\end{aligned}$$

Following central limit theorem, we have $\boldsymbol{S}_n^* \xrightarrow{d} N(0,\ \boldsymbol{V}_n^*)$. Therefore, we can obtain

$$\boldsymbol{T}_n^* \xrightarrow{d} \chi_K^2.$$

To obtain the desired result, we need to show

$$\boldsymbol{V}_n = \boldsymbol{V}_n^* + o_p(1), \tag{15}$$

and

$$\boldsymbol{S}_n = \boldsymbol{S}_n^* + o_p(1), \tag{16}$$

It is easy to show that (15) holds since by Theorem 1,

$$\|\hat{\boldsymbol{\eta}}_k(u_0) - \boldsymbol{\eta}_{k,0}\| = O_p(n^{-1/2}). \tag{17}$$

For (16), it is sufficient to show that $S_{n,k} = S_{n,k}^* + o_p(1)$ for any $1 \le k \le K$. Denote $S_{n,k}(\boldsymbol{\eta}_k) = n^{-1/2}\sum_{i=1}^n p_i^*(\tau_k; \boldsymbol{\eta}_k, u_0)\psi_{\tau_k}\{Y_i - Q_Y(\tau_k; \boldsymbol{\eta}_k, u_0|\mathbf{W}_i)\}$, because we have

$$\begin{aligned} S_{n,k}(\boldsymbol{\eta}_k) - S_{n,k}^* &= n^{-1/2}\sum_{i=1}^n p_i^*(\tau_k; \boldsymbol{\eta}_k, u_0)\psi_{\tau_k}\{Y_i - Q_Y(\tau_k; \boldsymbol{\eta}_k, u_0)\} \\ &\quad -n^{-1/2}\sum_{i=1}^n p_i^*(\tau_k; \boldsymbol{\eta}_{k,0}, u_0)\psi_{\tau_k}\{Y_i - Q_Y(\tau_k; \boldsymbol{\eta}_{k,0}, u_0)\}. \end{aligned}$$

Because $E(S_{n,k}^*) = 0$, following He and Shao (2000), we have

$$\sup_{\|\boldsymbol{\eta}_k - \boldsymbol{\eta}_{k,0}\| \le C_4 n^{-1/2}} \|S_{n,k}(\boldsymbol{\eta}_k) - S_{n,k}^* - E\{S_{n,k}(\boldsymbol{\eta}_k)\}\| = o_p(1), \tag{18}$$

where C_4 is some positive constant. For any $\boldsymbol{\eta}_k$ such that $\|\boldsymbol{\eta}_k - \boldsymbol{\eta}_{k,0}\| \le C_4 n^{-1/2}$, by the Taylor expansion, we get

$$\begin{aligned} E\{S_{n,k}(\boldsymbol{\eta}_k)\} &= n^{-1/2}\sum_{i=1}^n E(p_i^*(\tau_k; \boldsymbol{\eta}_k, u_0)[\tau_k - F_i\{Q_Y(\tau_k; \boldsymbol{\eta}_k, u_0)|\mathbf{W}_i\}]) \\ &= n^{-1/2}\sum_{i=1}^n E(p_i^*(\tau_k; \boldsymbol{\eta}_k, u_0)\,[-f_i\{Q_Y(\tau_k; \boldsymbol{\eta}_{k,0}, u_0)|\mathbf{W}_i\}\mathbf{m}^T(\mathbf{W}_i, u_0)(\boldsymbol{\eta}_k - \boldsymbol{\eta}_{k,0}) \\ &\quad - f_i^{'}\{Q_Y(\tau_k; \boldsymbol{\eta}_{k,0}, u_0)\}\{\mathbf{m}^T(\mathbf{W}_i, u_0)(\boldsymbol{\eta}_k - \boldsymbol{\eta}_{k,0})\}^2 + o_p(\|\boldsymbol{\eta}_k - \boldsymbol{\eta}_{k,0}\|^2)]) \\ &= -n^{-1/2}\sum_{i=1}^n E(\boldsymbol{p}_i^*(\tau_k; \boldsymbol{\eta}_k, u_0) f_i^{'}\{Q_Y(\tau_k; \boldsymbol{\eta}_{k,0}, u_0)|\mathbf{W}_i\} \\ &\quad [\{\mathbf{m}^T(\mathbf{W}_i, u_0)(\boldsymbol{\eta}_k - \boldsymbol{\eta}_{k,0})\}^2 + o(1)]) \\ &= o(1), \end{aligned} \tag{19}$$

where the third equality follows from the orthogonalization between $\boldsymbol{p}_i^*(\tau_k; \boldsymbol{\eta}_k, u_0)$ and $\mathbf{m}(\mathbf{W}_i,\ u_0)$, the fourth equality is based on $\|n^{-1}\sum_{i=1}^n E\boldsymbol{p}_i^*(\tau_k; \boldsymbol{\eta}_k, u_0)\| \le \|n^{-1}\sum_{i=1}^n E\boldsymbol{p}_i(\tau_k; \boldsymbol{\eta}_k, u_0)\| \le o(1)$. Let $\boldsymbol{\eta}_k = \hat{\boldsymbol{\eta}}_k(u_0)$, note that $S_{n,k} = S_{n,k}\{\hat{\boldsymbol{\eta}}_k(u_0)\}$. By combining (16) with (19), we obtain (16). This completes the proof of Theorem 2.
□

References

Bondell, H., Reich, B., Wang, J. (2010). Noncrossing quantile regression curve estimation. *Biometrika*, *97*(4), 825–838.

Bunker, C., Ukoli, F., Matthews, K., Kriska, A., Huston, S., Kuller, L. (1995). Weight threshold and blood pressure in a lean black population. *Hypertension*, *26*(4), 616–623.

Chan, K., Tsay, R. (1998). Limiting properties of the least squares estimator of a continuous threshold autoregressive model. *Biometrika*, *85*(2), 413–426.

Chappell, R. (1989). Fitting bent lines to data, with application to allometry. *Journal of Theoretical Biology*, *138*(2), 235–256.

Chen, L., Wei, Y. (2005). Computational issues for quantile regression. *Sankhyā*, 67(2):399–417.

Chernozhukov, V., Fernández-Val, I., Galichon, A. (2010). Quantile and probability curves without crossing. *Econometrica*, *78*(3), 1093–1125.

Dette, H., Volgushev, S. (2008). Non-crossing non-parametric estimates of quantile curves. *Journal of the Royal Statistical Society*: *Series B (Statistical Methodology)*, *70*(3), 609–627.

Feder, P. (1975). On asymptotic distribution theory in segmented regression problems identified case. *The Annals of Statistics*, *3*(1), 49–83.

Fiteni, I. (2004). τ-estimators of regression models with structural change of unknown location. *Journal of Econometrics*, *119*(1), 19–44.

Freedman, D. (1981). Bootstrapping regression models. *The Annals of Statistics*, *9*(6), 1218–1228.

Galvao, A. F., Kato, K., Montes-Rojas, G., Olmo, J. (2014). Testing linearity against threshold effects: uniform inference in quantile regression. *Annals of the Institute of Statistical Mathematics*, *66*(2), 413–439.

Gutenbrunner, C., Jurě, J., Koenker, R., Portnoy, S. (1993). Tests of linear hypotheses based on regression rank scores. *Journal of Nonparametric Statistics*, *2*(4), 307–331.

Hall, P., Sheather, S. (1988). On the distribution of a studentized quantile. *Journal of the Royal Statistical Society*: *Series B (Statistical Methodology)*, *50*(3), 381–391.

He, X. (1997). Quantile curves without crossing. *The American Statistician*, *51*(2), 186–192.

He, X., Shao, Q. (1996). A general Bahadur representation of M-estimators and its application to linear regression with nonstochastic designs. *The Annals of Statistics*, *24*(6), 2608–2630.

He, X., Shao, Q. (2000). On parameters of increasing dimensions. *Journal of Multivariate Analysis*, *73*(1), 120–135.

Hendricks, W., Koenker, R. (1992). Hierarchical spline models for conditional quantiles and the demand for electricity. *Journal of the American Statistical Association*, *87*(417), 58–68.

Jiang, L., Wang, H., Bondell, H. (2013). Interquantile shrinkage in regression models. *Journal of Computational and Graphical Statistics*, *22*(4), 970–986.

Jiang, X., Jiang, J., Song, X. (2012). Oracle model selection for nonlinear models based on weighted composite quantile regression. *Statistica Sinica*, *22*(4), 1479–1506.

Kai, B., Li, R., Zou, H. (2010). Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression. *Journal of the Royal Statistical Society*: *Series B (Statistical Methodology)*, *72*(1), 49–69.

Kai, B., Li, R., Zou, H. (2011). New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models. *The Annals of Statistics*, *39*(1), 305–332.

Kaufman, J., Asuzu, M., Mufunda, J., Forrester, T., Wilks, R., Luke, A., et al. (1997). Relationship between blood pressure and body mass index in lean populations. *Hypertension*, *30*(6), 1511–1516.

Kerry, S., Micah, F., Plange-Rhule, J., Eastwood, J., Cappuccio, F. (2005). Blood pressure and body mass index in lean rural and semi-urban subjects in West Africa. *Journal of Hypertension*, *23*(9), 1645–1651.

Kocherginsky, M., He, X., Mu, Y. (2005). Practical confidence intervals for regression quantiles. *Journal of Computational and Graphical Statistics*, *14*(1), 41–55.

Koenker, R. (1984). A note on l-estimates for linear models. *Statistics & Probability Letters*, *2*(6):323–325.

Koenker, R. (1994). Confidence intervals for regression quantiles. *Proceedings of the 5th Prague symposium on asymptotic statistics* (pp. 349–359). New York: Springer.

Koenker, R. (2005). *Quantile regression*. Cambridge: Cambridge University Press.

Koenker, R., Bassett, G. (1978). Regression quantiles. *Econometrica*, *46*(1), 33–50.

Kosorok, M., Song, R. (2007). Inference under right censoring for transformation models with a change-point based on a covariate threshold. *The Annals of Statistics*, *35*(3), 957–989.

Lee, S., Seo, M., Shin, Y. (2011). Testing for threshold effects in regression models. *Journal of the American Statistical Association*, *106*(493), 220–231.

Li, C., Wei, Y., Chappell, R., He, X. (2011). Bent line quantile regression with application to an allometric study of land mammals' speed and mass. *Biometrics*, *67*(1), 242–249.

Liu, J., Wu, S., Zidek, J. (1997). On segmented multivariate regression. *Statistica Sinica*, *7*(2), 97–525.

Liu, Z., Qian, L. (2010). Changepoint estimation in a segmented linear regression via empirical likelihood. *Communications in Statistics—Simulation and Computation*, *39*(1), 85–100.

Muggeo, V. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine*, *22*(19), 3055–3071.

Oka, T., Qu, Z. (2011). Estimating structural changes in regression quantiles. *Journal of Econometrics*, *162*(2), 248–267.

Pastor, R., Guallar, E. (1998). Use of two-segmented logistic regression to estimate change-points in epidemiologic studies. *American Journal of Epidemiology*, *148*(7), 631–642.

Qu, Z. (2008). Testing for structural change in regression quantiles. *Journal of Econometrics*, *146*(1), 170–184.

Robinson, D. (1964). Estimates for the points of intersection of two polynomial regressions. *Journal of the American Statistical Association*, *59*(305), 214–224.

Su, L., Xiao, Z. (2008). Testing for parameter stability in quantile regression models. *Statistics & Probability Letters*, *78*(16):2768–2775.

Vieth, E. (1989). Fitting piecewise linear regression functions to biological responses. *Journal of the American Statistical Association*, *67*(1), 390–396.

Wang, J., He, X. (2007). Detecting differential expressions in genechip microarray studies: A quantile approach. *Journal of the American Statistical Association*, *102*(477), 104–112.

Wang, J., Zhu, Z., Zhou, J. (2009). Quantile regression in partially linear varying coefficient models. *The Annals of Statistics*, *37*(6B), 3841–3866.

Zhang, L., Wang, J., Zhu, Z. (2014). Testing for change points due to a covariate threshold in quantile regression. *Statistica Sinica*, *24*(4), 1859–1877.

Zou, H., Yuan, M. (2008). Composite quantile regression and the oracle model selection theory. *The Annals of Statistics*, *36*(3), 1108–1126.