Article

The Estimation of Bent Line Expectile Regression Model Based on a Smoothing Technique

Jie Liu [1], Jiaqing Chen [1,*] and Yangxin Huang [2]

[1] Department of Statistics, Wuhan University of Technology, Wuhan 430070, China; 238645@whut.edu.cn
[2] Department of Epidemiology and Biostatistics, University of South Florida, Tampa, FL 33620, USA; yhuang@health.usf.edu
* Correspondence: jqchenwhut@163.com

Abstract: A bent line expectile regression model can describe the effect of a covariate on the response variable with two different straight lines, which intersect at an unknown change-point. Due to the existence of the change-point, the objective function of the model is not differentiable with respect to the change-point, so it cannot be solved by the method of the traditional linear expectile regression model. For this model, a new estimation method is proposed by a smoothing technique, that is, using Gaussian kernel function to approximate the indicator function in the objective function. It can not only estimate the regression coefficients and change-point location simultaneously, but also have better estimation effect, which compensates for the insufficiency of the previous estimation methods. Under the given regularity conditions, the theoretical proofs of the consistency and asymptotic normality of the proposed estimators are derived. There are two parts of numerical simulations in this paper. Simulation 1 discusses various error distributions at different expectile levels under different conditions, the results show that the mean values of the biases of the estimation method in this paper, and other indicators, are very small, which indicates the robust property of the new method. Simulation 2 considers the symmetric and asymmetric bent lien expectile regression models, the results show that the estimated values of the estimation method in this paper are similar to the true values, which indicates the estimation effect and large sample performance of the proposed method are excellent. In the application research, the method in this paper is applied to the Arctic annual average temperature data and the Nile annual average flow data. The research shows that the standard errors of the estimation method in this paper are very similar to 0, indicating that the parameter estimation accuracy of the new method is very high, and the location of the change-point can be accurately estimated, which further confirms that the new method is effective and feasible.

Keywords: bent line expectile model; change-point; smoothing technique; estimation

1. Introduction

When performing regression analysis on data in the fields of financial economics [1,2], biomedicine [3,4] and environmental science [5,6], it is inevitable to encounter changes in the data structure. The point corresponding to the change in the data structure is called a change-point [7]. Establishing piecewise linear regression model to study the change-point problem is one of the hot issues in recent years. At present, there have been a large number of literature studies on the piecewise linear regression model, which has been widely used in various fields [8–11]. According to whether the regression function of the piecewise linear regression model is continuous at the change-point, Bhattacharya [12] divided the model into two categories. One is that the regression function of the model is continuous at the change-point, which is called continuous threshold regression model, and the corresponding change-point is called continuous change-point. On the contrary, the other is called discontinuous piecewise linear regression model, and the corresponding change-point is called discontinuous change-point. This paper is to study the continuous threshold regression model.

In the aspect of change-point detection, Chan [13] extended the log-likelihood ratio test statistics to the piecewise autoregressive model to detect the existence of the change-point. Liu and Qian [14] introduced an empirical likelihood ratio test statistics for continuous change-point in the regression model. Lee et al. [15] proposed general statistics for change-point detection in the regression model, namely sup-likelihood ratio test statistics, which shows that the asymptotic distribution of the test statistics under null assumption is non-standard. In addition, some scholars have studied the change-point detection problem of non-random covariates, mainly focusing on the analysis of time series data, specific references can be referred to Hansen [16] and Cho and White [17], among others. It should be emphasized that the main research goal of this paper is to propose a new estimation method for the continuous threshold expectile regression model, rather than detecting the existence of the change-point. Therefore, this paper is based on the assumption of the existence of the change-point.

In some practical applications, except the conditional expectation of the response variable, we want to have more complete information on the response variable. The expectile regression model proposed by Newey and Powell [18] can not only provide complete information on the response variable through tail expectation, but also have loose application conditions and no strict assumptions for error terms. More importantly, its loss function is differentiable, so it is convenient to calculate. It is a good supplement and expansion to the traditional mean regression model and quantile regression model. Zhang and Li [19] firstly introduced a continuous threshold expectile regression model, also known as bent line expectile regression model, and proposed a formal test for the existence of the change-point at given expectile levels. Similar to Lerman [20], Zhang and Li [19] proposed a grid search method to estimate the regression coefficients and change-point parameter, respectively. The main idea of the grid search method is to search over all possible values under the domain of the change-point and then to select the best one. Nevertheless, the grid search method estimates the regression coefficients and change-point location separately. Not only that, it requires a finer grid to identify the change-point location more precisely, which could result in expensive computation. Furthermore, it is not realistic for this method by assuming that the change-point only occurs at discrete grid points. Motivated by these issues, Zhou and Zhang [21] extended the linearization technique proposed by Yan et al. [22] to the bent line expectile regression model and constructed the interval estimation of the estimators through the theory of the standard linear expectile regression model and the delta technique. Nonetheless, estimation obtained by using the linearization technique in general may underestimate the change-point.

Smoothing technique is to use Gaussian kernel function to approximate the indicator function in the objective function, so as to achieve the purpose of smoothing the change-point locally. Then the objective function is continuous and differentiable with respect to all parameters after being smoothed, so we can easily estimate all parameters in the model. This paper is to propose a new estimation method for the bent line expectile regression model through a smoothing technique, which can simultaneously compensate for the insufficiency of the grid search method of Zhang and Li [19] and the linearization method of Zhou and Zhang [21]. In addition, the theoretical proofs of the consistency and asymptotic normality of the proposed estimators are derived. The numerical simulations under the different conditions show the good performance of the new estimation method proposed in this paper. Finally, the research based on the Arctic annual average temperature data and the Nile annual average flow data show that the new method is effective and feasible.

The remainder of this paper is structured as follows. Section 2 introduces the bent line expectile regression model and proposes a new estimation method for the model based on a smoothing technique. At the same time, we also establish the asymptotic properties of the estimators. In Section 3, simulation studies are conducted to investigate the finite sample performance of the estimation method proposed in this paper. Section 4 applies the model and the estimation method to analyze two real data. Section 5 concludes the paper.

2. Methodology

In this section, we first introduce the bent line expectile regression model, and then point out the insufficiency of the grid search method and linearization method to introduce the smoothing method proposed in this paper. Finally, we prove the consistency and asymptotic normality of the proposed estimators obtained by this new method.

2.1. Model Introduction

Let $(Y_i,\ X_i, \mathbf{Z}_i)$, $i = 1, \ldots, n$ be the independent identically distributed random samples from the population $(Y, X, \mathbf{Z})$. Given $0 < \tau < 1$, Zhang and Li [19] proposed a bent line expectile regression model:

$$Q_\tau(Y|X, \mathbf{Z}) = \beta_0 + \beta_1 X + \beta_2(X - t)_+ + \gamma^\mathsf{T} \mathbf{Z} \tag{1}$$

where Y is the response variable, X is the scalar covariate with a change-point, $\mathbf{Z}$ is a $p \times 1$-dimensional vector of linear covariates with constant slopes, $Q_\tau(Y|X, \mathbf{Z})$ is the τth conditional expectile of the response variable Y given X and $\mathbf{Z}$, $u_+ = uI\ (u > 0)$ and $I(\cdot)$ is the indicator function, t is an unknown change-point, $\xi = (\beta_0, \beta_1, \beta_2, \gamma^\mathsf{T})^\mathsf{T}$ are the regression coefficients with the exception of t. In model (1), the linear expectile regression model $Q_\tau(Y|X, \mathbf{Z})$ is continuous on the covariate X at the change-point location t, but X has segmented effects on the expectile of Y through the unknown change-point t. More specifically, the slope of X is β_1 below the change-point t, whereas it is $\beta_1 + \beta_2$ above that point t. For the identifiability of the change-point, it is commonly assumed that $\beta_2 \neq 0$.

To estimate all the regression parameters of the model $\theta = (\xi^\mathsf{T}, t)^\mathsf{T}$ at given expectile levels $\tau \in (0,1)$, an estimator for θ can be obtained by minimizing the following objective function:

$$l_{n,\tau}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \rho_\tau \big\{ Y_i - \beta_0 - \beta_1 X_i - \beta_2(X_i - t)_+ - \gamma^\mathsf{T} \mathbf{Z}_i \big\} \tag{2}$$

where $\rho_\tau(u) = u^2|\tau - I(u < 0)|$ is the asymmetric least square (ALS) loss function.

2.2. The Proposed Method

However, due to the existence of the change-point t, the objective function (2) is not differentiable with respect to t, which leads to the non-smoothness problem of the objective function, so it cannot be solved by the method of the traditional linear expectile regression model. In this regard, Zhang and Li [19] proposed the grid search method to circumvent the non-smoothness problem of the objective function. The main idea of the grid search method is to fix the location of the change-point on the grid in advance, then estimate the regression coefficients, and finally use the grid search algorithm to select the optimal point on the grid as the estimation of the change-point parameter. To be specific, for a fixed t, the parameter vector ξ can be estimated by,

$$\hat{\xi}(t) = \arg\min_{\xi} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau \{ Y_i - \mathbf{U}_i(t)\xi \}$$

where $\mathbf{U}_i(t) = \big(1, X_i, (X_i - t)_+, \mathbf{Z}_i^\mathsf{T}\big)$. Then, the change-point t can be estimated by

$$\hat{t} = \arg\min_{t} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau \big\{ Y_i - \mathbf{U}_i(t)\hat{\xi}(t) \big\}$$

Therefore, the estimator of θ is $\hat{\theta} = \big(\hat{\xi}(t)^\mathsf{T}, \hat{t} \big)^\mathsf{T}$.

Although the grid search method has become the mainstream estimation method of the change-point model, it has its limitations: the grid search method estimates the regression coefficients and change-point location separately; it requires a finer grid to identify the

change-point location more precisely, which could result in expensive computation; it is not realistic for this method by assuming that the change-point only occurs at discrete grid points. In order to overcome these defects of the grid search method, Zhou and Zhang [21] extended the linearization technique proposed by Yan et al. [22] to the bent line expectile regression model, converted the bent line expectile regression model into a standard linear expectile regression model, and then obtained the estimations of the regression coefficients and change-point parameter in the model (1) simultaneously through an iterative algorithm. The specific approach is to approximate the non-linear term by a first-order Taylor expansion around an initial value $t^{(0)}$,

$$(X_i - t)_+ \approx \left(X_i - t^{(0)}\right)_+ + (-1)I(X_i > t^{(0)})\left(t - t^{(0)}\right)$$

where $(-1)I(X_i > t^{(0)})$ is the first-order derivative of $(X_i - t)_+$ at $t^{(0)}$.

Although the estimation method proposed by Zhou and Zhang [21] makes up for the shortcomings of the grid search method, their method may also have a certain drawback, that is, estimation obtained by using the linearization technique in general may underestimate the change-point. Therefore, our next research content is to propose a better estimation method for the bent line expectile regression model, which can make up for the shortcomings of the grid search method and linearization method simultaneously.

Gaussian kernel function is not only simple in mathematical form and differentiable, that can make calculation easier, but also smooth, which can make the final model smoother. Therefore, we consider to use Gaussian kernel function $\Phi\left(\frac{X-t}{h}\right)$ to approximate the indicator function $I(X > t)$, where $\Phi(\cdot)$ is a standard normal distribution function, $h > 0$ is a bandwidth and $h \to 0$ when n tends to infinity. Consequently, the objective function (2) can be approximated by,

$$\ell(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\rho_\tau\left\{Y_i - \beta_0 - \beta_1 X_i - \beta_2(X_i - t)\Phi\left(\frac{X_i - t}{h}\right) - \boldsymbol{\gamma}^\mathsf{T}\mathbf{Z}_i\right\} \tag{3}$$

Note that the function in the bracket of (3) is continuous and differentiable with respect to all parameters $\boldsymbol{\theta}$, including the change-point t. Therefore, we can easily estimate the change-point parameter t and the regression parameters $\boldsymbol{\xi}$ in the model by minimizing the smoothed objective function.

2.3. Asymptotic Properties

To simplify notations, let,

$$g(\boldsymbol{w}_i, \boldsymbol{\theta}) = \beta_0 + \beta_1 X_i + \beta_2(X_i - t)\Phi\left(\frac{X_i - t}{h}\right) + \boldsymbol{\gamma}^\mathsf{T}\mathbf{Z}_i$$

where $\boldsymbol{w}_i = \left(1, X_i, \mathbf{Z}_i^\top\right)^\top$, $\boldsymbol{\theta} = \left(\beta_0, \beta_1, \beta_2, \boldsymbol{\gamma}^\top, t\right)^\top$. Then the parameters $\boldsymbol{\theta}$ can be estimated by minimizing the following objective function:

$$\ell_\tau(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\rho_\tau\{Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta})\} \tag{4}$$

that is,

$$\hat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta}}\ell_\tau(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\theta}}\frac{1}{n}\sum_{i=1}^{n}\rho_\tau\{Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta})\} \tag{5}$$

To obtain the asymptotic properties of the estimator $\hat{\boldsymbol{\theta}}_n$, we introduce more notations:

$$q(\boldsymbol{w}_i, \boldsymbol{\theta}) = \frac{\partial g(\boldsymbol{w}_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$$

$$= \left[1, X_i, (X_i - t)\Phi\left(\frac{X_i-t}{h}\right), \boldsymbol{Z}_i^\top, -\beta_2\Phi\left(\frac{X_i-t}{h}\right) - \frac{\beta_2(X_i-t)}{h}\,\Phi\prime\left(\frac{X_i-t}{h}\right) \right]^\top$$

$$C_h(\boldsymbol{\theta}_0) = \frac{[(1-2\tau)F_i(g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)) + \tau]}{n} \sum_{i=1}^{n} q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)q^\top(\boldsymbol{w}_i, \boldsymbol{\theta}_0)$$

$$D_h(\boldsymbol{\theta}_0) = -\frac{1}{n} \sum_{i=1}^{n} \left\{ q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)q^\top(\boldsymbol{w}_i, \boldsymbol{\theta}_0)[(1-2\tau)F_i(g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)) + \tau] \right\}$$

where $\boldsymbol{\theta}_0$ is the true parameter, $F_i(g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))$ is the distribution function of Y_i.

Obviously, minimizing the objective function (4) is equivalent to solving the equation,

$$\sum_{i=1}^{n} \psi_\tau\{Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta})\}q(\boldsymbol{w}_i, \boldsymbol{\theta}) = 0 \tag{6}$$

where $\psi_\tau(u) = u|\tau - I(u < 0)|$ is the kernel of the first-order derivative of the function $\rho_\tau(u)$. Solving the equation yields the estimators of the parameters.

The following theorem gives the asymptotic properties of the estimator $\hat{\boldsymbol{\theta}}_n$.

Theorem 1 (Asymptotic Properties). *Let $\boldsymbol{\theta}_0$ be the true parameter, and $\hat{\boldsymbol{\theta}}_n$ is the estimator of the estimation method in this paper. Under the regularity conditions in the* Appendix A, *we have $\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 = O_p\left(n^{-1/2}\right)$, and $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)$ is asymptotically normally distributed with mean zero and the covariance matrix of $\boldsymbol{\Sigma} = \tau(1-\tau)D_h^{-1}(\boldsymbol{\theta}_0)C_h(\boldsymbol{\theta}_0)D_h^{-\top}(\boldsymbol{\theta}_0)$.*

The matrices $C_h(\boldsymbol{\theta}_0)$ and $D_h(\boldsymbol{\theta}_0)$ in the covariance matrix can be estimated by the following formulas:

$$\hat{C}_h(\hat{\boldsymbol{\theta}}_n) = \frac{1}{n} \sum_{i=1}^{n} q(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n)q^\top(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n)$$

$$\hat{D}_h(\hat{\boldsymbol{\theta}}_n) = -\frac{1}{n} \sum_{i=1}^{n} \left\{ \left[(1-2\tau)\hat{F}_i(g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)) + \tau\right] q(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n)q^\top(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) \right\}$$

where $\hat{F}_i(g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))$ is the empirical distribution function of Y_i.

Note that the estimation of $\boldsymbol{\theta} = \left(\beta_0, \beta_1, \beta_2, \boldsymbol{\gamma}^\top, t\right)^\top$ is based on the choice of bandwidth h. Therefore, the choice of bandwidth is crucial. We use the cross-validation (*CV*) criterion to choose the bandwidth h. Specifically, let,

$$CV(h) = \sum_{i=1}^{n} \rho_\tau\{Y_i - \hat{Q}_{Y_{-i}}(\tau, \boldsymbol{\theta}X_i, \boldsymbol{Z}_i)\}$$

where $\hat{Q}_{Y_{-i}}$ is the estimator of the model after removing the ith observation $(Y_i, X_i, \boldsymbol{Z}_i)$. The bandwidth h that minimizes the value of $CV(h)$ is optimal.

3. Simulation Studies

There are two parts of numerical simulations in this section to verify the finite sample performance of the estimation method in this paper. The first simulation is to evaluate the robust properties of the proposed estimators with three different types of error. Another simulation considers the symmetric and asymmetric bent line expectile regression models.

3.1. Simulation 1

Considering that the independent identically distribution assumption can greatly reduce the case of individual cases in the sample, which makes the sample data more

representative in general, we consider the scenario of independent identically distribution. However, in many practical studies, the assumption of independent identically distribution is usually not satisfied, so corresponding to this, we consider the scenario of heteroscedasticity. On the other hand, there are many possible cases for the error term of the model. We discuss the two opposite cases that the error term obeys normal distribution and non-normal distribution. Furthermore, we study the mixed distribution which contains both normal distribution and non-normal distribution.

The simulation data in this section come from the following two scenarios:

(I) Independent identically distribution (IID): $Y = \beta_0 + \beta_1 X + \beta_2 (X - t)_+ + \gamma Z + e$
(II) Heteroscedasticity: $Y = \beta_0 + \beta_1 X + \beta_2 (X - t)_+ + \gamma Z + (1 + 0.3X)e$

where $X \sim U(-2, 5)$, $Z \sim B(1, 0.5)$. The τth expectile of the random error term e is 0, that is, $e = \widetilde{e} - Q_\tau(\widetilde{e})$, where $Q_\tau(\widetilde{e})$ is the τth expectile of $\widetilde{e}$.

For each scenario, we consider the following three different error cases:

(1) a standard normal distribution: $\widetilde{e} \sim N(0, 1)$
(2) a t-distribution with three degrees of freedom: $\widetilde{e} \sim t_3$
(3) a contaminated standard normal distribution: $\widetilde{e} \sim 0.9N(0, 1) + 0.1t_3$

The regression coefficients are set as $(\beta_0, \beta_1, \beta_2, \gamma) = (1, 3, -2, 1)$ and the change-point is $t = 1.5$. For each case, we conduct 200 repetitions with sample sizes $n = 200$.

In order to compare the proposed method with that of Zhou and Zhang [21], we also consider the linearization method. Tables A1–A6 show the simulation results of the linearization method ("seg") and the proposed method in this paper ("proposed") at the different expectile levels $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$. It should be noted that the data in the tables represent biases (estimated value − true value), that is, *bias* $= \hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0$. It can be seen from the tables that the mean values ("Mean") of the biases of the proposed method in this paper are very similar to 0 at different expectile levels of various error distributions, and for some parameters, this indicator is even equal to 0 in some cases, for example, for the independent identically distribution model with $\widetilde{e} \sim t_3$, the mean value ("Mean") of the bias of γ is 0 when $\tau = 0.5$. In addition, this indicator of the method in this paper is far less than the linearization method. On the other hand, the sum of squared errors ("SSE") of the biases of the two estimation methods are very similar to 0, indicating that the fitting results of the two methods are so satisfying. Further, we find that this indicator of the method in this paper is only about half of that of the linearization method through comparison, which indicates that the method in this paper is more suitable for data fitting. In addition, the minimums ("Min") and the maximums ("Max") of the biases of the proposed method are similar to those of the linearization method, and the value ranges of the minimums ("Min") and the maximums ("Max") of the two methods are generally $(0, 1)$, just for the case of $\widetilde{e} \sim t_3$, there are more cases where the two indicators are greater than 1 in the two methods, but this situation is also normal. In addition, the upper quartiles ("IQ1"), the medians ("Median") and the lower quartiles ("IQ3") of the biases of the proposed method are slightly smaller than those of the linearization method, which is especially evident at the tail expectile levels (e.g., $\tau = 0.1, 0.9$). It shows that estimation obtained by using the linearization technique indeed underestimates the change-point. All in all, the estimation method in this paper is better than the linearization method.

3.2. Simulation 2

Considering that the location of the change-point may be different, we divide the sources of the simulation data into symmetric and asymmetric cases, which could ensure the rigor and comprehensiveness of the study.

The simulation data in this section come from the following two scenarios:

Symmetric scenario:

$$Y = 0 + \beta_1 X - 2\beta_1 (X - 5)_+ + c_1(X)e$$

Asymmetric scenario:

$$Y = 4.5 - \beta_1 + \beta_1 X - (\beta_1 + 0.5)(X - 2)_+ + c_2(X)e$$

where $c_1(X) = 0.5 + 0.1X$, $c_2(X) = 0.5 + 0.02X$, e is standard normal distribution and its τth expectile is 0.

In the symmetric scenario, the scalar covariate $X \sim U(0, 10)$, and the change-point is the median of X which is $t = 5$. Then corresponding to the model in this paper, we set $(\beta_0, \beta_1, \beta_2, t) = (0, \beta_1, -2\beta_1, 5)$. In the asymmetric scenario, the scalar covariate X is a mixture distribution that uniformly on $(0, 1)$ with probability 0.1 and uniformly on $(0, 10)$ with probability 0.9, that is, $X \sim 0.1U(0, 1) + 0.9U(0, 10)$, the change-point is $t = 2$. Then corresponding to the model in this paper, we set $(\beta_0, \beta_1, \beta_2, t) = (4.5 - \beta_1, \beta_1, -\beta_1 - 0.5, 2)$.

For each scenario, we set the parameter $\beta_1 = 0.5$, and we conduct 200 repetitions with sample sizes $n = 200$ by applying the linearization method ("seg") and our method ("proposed"). Tables A7 and A8 record the simulation results at different expectile levels $\tau = 0.1, 0.3, 0.5, 0.7, 0.9$. It is also worth noting that the data in the tables also represent biases. It can be seen from the tables that the results obtained by the method in this paper and the linearization method are very similar, and their indicators are relatively small, which also shows that the estimated values of both methods are similar to the true values. Not only that, we discover that the mean values ("Mean"), the sum of squared errors ("SSE"), the minimums ("Min"), the upper quartiles ("IQ1"), the medians ("Median"), the lower quartiles ("IQ3") and the maximums ("Max") of the biases of the two methods increase with τ, which indicates that the fitting effect is better while the expectile is smaller. In addition, when τ is small (e.g., $\tau = 0.1, 0.3$), these indicators of the method in this paper are slightly smaller than those of the linearization method. However, when τ is larger (e.g., $\tau = 0.7, 0.9$), these indicators of the proposed method are slightly larger than those of the linearization method. In general, the estimation effect and large sample performance of the proposed method are good.

4. Empirical Applications

There are two parts of practical applications in this section. We apply the proposed method in this paper to the Arctic annual average temperature data and the Nile annual average flow data separately to illustrate the practical utility of the new method.

4.1. Application 1

The latest evidence shows that the earth has been warming over the past 50 years, and this trend has not changed. Scientists from many countries have contributed to the study of global climate change, among which the warming of the Arctic climate is one of the important aspects of the study of global climate change.

In this section, we mainly focus on the annual average temperature data in the Arctic region (90 N–23.6 N) from 1940 to 2021. This dataset is available from NASA Goddard Institute for Space Studies (www.giss.nasa.gov (accessed on 15 March 2022)). Based on the scatter diagram provided by the institute, we can roughly learn that the annual average temperature in the Arctic region remained almost stable between 1940 and 1975, and then increased rapidly after 1975. Therefore, we guess the annual average temperature in the Arctic region changed in 1975. Therefore, we fit the following model to the Arctic annual average temperature data:

$$Q_\tau(Y|X) = \beta_0 + \beta_1 X + \beta_2 (X - t)_+$$

where Y is the annual average temperature in the Arctic region, X is the continuous time and $X_1 < X_2 < \ldots < X_n$, $n = 82$ is the sample size. We consider $\tau = 0.2, 0.3, 0.5, 0.7, 0.8$ these five different expectile levels. Table A9 summarizes the estimations ("Estimate") of the parameters and their standard errors ("SE") of the grid search method, linearization method and smoothing method proposed in this paper. It can be seen from the table that

the parameter estimation accuracy of the smoothing method proposed in this paper is exactly the same as that of the grid search method, the standard errors of all parameters except β_2 are equal to 0, and the standard error of β_2 is also very similar to 0, which indicates that the parameter estimation accuracy of the two methods are all very high. This is not difficult to explain, although the grid search method is complex and its computation is very large, as long as the fine grid is divided to a certain degree, the optimal value can be obtained, which also shows the good performance and practical utility of the smoothing method proposed in this paper. Moreover, we find that the parameter estimation accuracy of the smoothing method and the grid search method is slightly higher than that of the linearization method. On the other hand, as far as our dataset is concerned, the smoothing method and linearization method can save about a quarter of computation and time compared with the grid search method, which greatly improves the computational efficiency. Most importantly, we can intuitively see the change-point $t \approx 1975$, that is, the annual average temperature in the Arctic region changed significantly in 1975, which also verifies our conjecture.

4.2. Application 2

The Nile dataset is the internal dataset in R, which represents the measurements of the annual average flow (in unit 10^8 m^3) of the River Nile at Aswan between 1871 and 1970. In this section, we mainly focus on the annual average flow data of the Nile from 1871 to 1970. Based on the scatter diagram in Figure 1, we can roughly learn that the annual average flow of the River Nile showed a downward trend before 1914, then increased rapidly after 1914. In addition, it should be noted that the number 1 in the abscissa of the graph represents 1871, and so on.
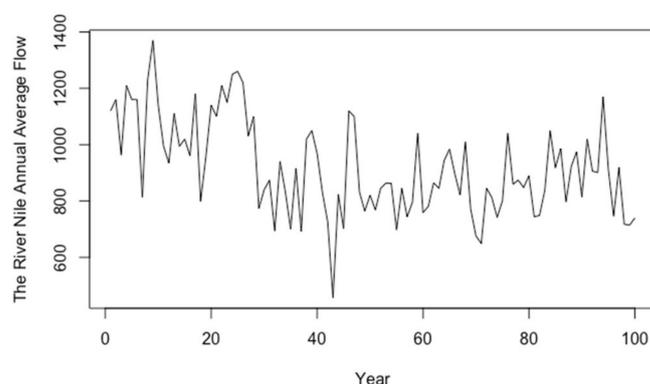
Figure 1. The scatter diagram of the annual average flow of the River Nile.

Therefore, we fit the following model to the Nile annual average flow data:

$$Q_\tau(Y|X) = \beta_0 + \beta_1 X + \beta_2 (X - t)_+$$

where Y is the annual average flow of the River Nile, X is the continuous time and $X_1 < X_2 < \ldots < X_n$, $n = 100$ is the sample size. We consider $\tau = 0.2, 0.3, 0.5, 0.7, 0.8$ these five different expectile levels. Table A10 summarizes the estimations ("Estimate") of the parameters and their standard errors ("SE") of the grid search method, linearization method and smoothing method proposed in this paper. It can be seen from the table that the parameter estimation accuracy of the smoothing method proposed in this paper is very similar to that of the grid search method, and both of them are very high. Not only that, we find this indicator of the proposed method and grid search method is slightly higher than that obtained by the linearization method. On the other hand, as far as our dataset is concerned, compared with the grid search method, the smoothing method and linearization method have smaller computation and less operation time. In addition, we can intuitively see $t \approx 1914$ obtained by the proposed method when $\tau = 0.2, 0.3, 0.5$, that is,

the annual average flow of the River Nile changed significantly in 1914. It also shows that $t \approx 1924$ when $\tau = 0.7, 0.8$, and this situation is also acceptable.

5. Conclusions

In this paper, we study the problem of statistical inference of the change-point and other regression coefficients of the bent line expectile regression model. Due to the existence of the change-point, the objective function of the model is not differentiable with respect to the change-point, which brings a huge challenge to our calculation. In order to compensate for the insufficiency of the previous estimation methods, we propose a new estimation method by a smoothing technique, that is, using Gaussian kernel function to approximate the indicator function in the objective function, which can not only estimate the regression coefficients and change-point location simultaneously, greatly reduce the computation, but also improve the accuracy of the parameter estimation. At the same time, we give the large sample properties of the estimation method in this paper, the estimator $\hat{\boldsymbol{\theta}}_n$ converges to the true parameter $\boldsymbol{\theta}_0$ at a speed of $n^{-1/2}$ in probability, and the bias $\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0$ is asymptotically normally distributed. Through the numerical simulation analysis, we find that all the indicators of the biases of the parameter estimations are almost equal to 0, which shows that the new method proposed in this paper has a very good estimation effect and verifies the large sample performance of the new method. Finally, we also apply the model and the method in this paper to the analysis of two actual data and find that the estimated values of the parameters obtained by the new method proposed in this paper are almost equal to the true values, which shows the practical utility of the new method. It not only improves the accuracy of the parameter estimation, but also greatly improves the computational efficiency compared with the grid search method.

The model studied in this paper is the continuous threshold expectile regression model with a single change-point, but in practical applications, it is very meaningful to consider a continuous threshold expectile regression model with multiple change-points. In this regard, we can consider extending the current research work to the multiple change-points situation, which is believed to be a very interesting topic.

Appendix A. Proof of Theorem 1

Without loss of generality, the asymptotic properties of the proposed estimators are derived based on the following assumptions:

A1 There is a constant $\epsilon > 0$ that holds $\lim\limits_{n \to \infty} P(|X_i - t| < \epsilon) = 0$.

A2 For all $i = 1, \ldots, n$, $F_\tau \{ Q_{Y_i}(\tau, \boldsymbol{\theta} \mid \boldsymbol{w}_i) \}$ is the conditional expectile distribution function of the response variable Y_i given $\boldsymbol{w}_i$ and has a bounded and continuous expectile density function $f_\tau \{ Q_{Y_i}(\tau, \boldsymbol{\theta} \mid \boldsymbol{w}_i) \}$.

A3 The density function $p_i(x)$ of X_i is continuous on the bounded compact set $[M_1, M_2]$.

A4 $E\left(\parallel \boldsymbol{Z}_i \parallel^3\right) < \infty$, where $\parallel \cdot \parallel$ is the Euclidean norm of any vector.

A5 The bandwidth $h \to 0$ as $n \to \infty$.

A6 There are two positive definite matrices $C_{0,\tau}$ and $D_{0,\tau}$ that hold $\lim\limits_{n \to \infty} C_{n,\tau} = C_{0,\tau}$ and $\lim\limits_{n \to \infty} D_{n,\tau} = D_{0,\tau}$.

Lemma A1 (Consistency). *Under the conditions A1–A5, we let $\hat{\boldsymbol{\theta}}_n$ be the estimator of the minimized objective function $\ell_\tau(\boldsymbol{\theta})$ and $\boldsymbol{\theta}_0$ be the true parameter. Then $\hat{\boldsymbol{\theta}}_n$ converges to $\boldsymbol{\theta}_0$ in probability, that is to say, it is established $\hat{\boldsymbol{\theta}}_n \overset{p}{\to} \boldsymbol{\theta}_0$.*

Proof of Lemma A1. To prove the consistency of the estimators, we firstly have to show that,

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} |l(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta})| \overset{p}{\to} 0$$

Based on assumption A1, for a certain $\epsilon > 0$, we have $\lim\limits_{n \to \infty} P(|X_i - t| < \epsilon) = 0$. It is easy to show that for any $\boldsymbol{\xi} > 0$, there exists $N > 0$ that satisfies $P(|X_i - t| < \epsilon) < \boldsymbol{\xi}$ when $i > N$. In addition, based on the definition of the function $\Phi(\cdot)$, if satisfied $|X_i - t| \geq \epsilon$, then there is,

$$\Phi\left(\frac{X_i - t}{h}\right) = I(X_i > t) + o_i(h), h \to 0.$$

Let $v(\boldsymbol{w}_i, \boldsymbol{\theta}) = \beta_0 + \beta_1 X_i + \beta_2(X_i - t)I(X_i > t) + \boldsymbol{\gamma}^\mathsf{T} \boldsymbol{Z}_i$. Notice,

$$g(\boldsymbol{w}_i, \boldsymbol{\theta}) = v(\boldsymbol{w}_i, \boldsymbol{\theta}) + \beta_2(X_i - t)\left(\Phi\left(\frac{X_i - t}{h}\right) - I(X_i > t)\right)$$
$$= v(\boldsymbol{w}_i, \boldsymbol{\theta}) + r(\boldsymbol{w}_i, \boldsymbol{\theta})$$

where $r(\boldsymbol{w}_i, \boldsymbol{\theta}) = \beta_2(X_i - t)\left(\Phi\left(\frac{X_i - t}{h}\right) - I(X_i > t)\right)$.

Then we can easily acquire,

$$l(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta})$$
$$= \frac{1}{n}\sum_{i=1}^n \{\rho_\tau(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta})) - \rho_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))\}$$
$$= \frac{1}{n}\sum_{i=1}^n \Big[(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 |\tau - \mathrm{I}(Y_i < v(\boldsymbol{w}_i, \boldsymbol{\theta}))|$$
$$- (Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 |\tau - \mathrm{I}(Y_i < g(\boldsymbol{w}_i, \boldsymbol{\theta}))| \Big]$$
$$= \frac{1}{n}\sum_{i=1}^n \Big\{ (Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 (|\tau - \mathrm{I}(Y_i < v(\boldsymbol{w}_i, \boldsymbol{\theta}))| - |\tau - \mathrm{I}(Y_i < g(\boldsymbol{w}_i, \boldsymbol{\theta}))|)$$
$$+ \Big[(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 - (Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \Big] |\tau - \mathrm{I}(Y_i < g(\boldsymbol{w}_i, \boldsymbol{\theta}))| \Big\}$$
$$= \frac{1}{n}\sum_{i=1}^n \Big[(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 (1 - 2\tau) I(|Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta})| \langle |r(\boldsymbol{w}_i, \boldsymbol{\theta})|) * \mathrm{sgn}(-r(\boldsymbol{w}_i, \boldsymbol{\theta}))$$
$$+ (2Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}) - g(\boldsymbol{w}_i, \boldsymbol{\theta}))r(\boldsymbol{w}_i, \boldsymbol{\theta}) |\tau - \mathrm{I}(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}) < r(\boldsymbol{w}_i, \boldsymbol{\theta}))| \Big]$$
$$= \frac{1}{n}\sum_{i=1}^n \Big[(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 (1 - 2\tau) \mathrm{I}\{|Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta})| < |r(\boldsymbol{w}_i, \boldsymbol{\theta})|\} \mathrm{sgn}\{-r(\boldsymbol{w}_i, \boldsymbol{\theta})\}$$
$$+ 2(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 r(\boldsymbol{w}_i, \boldsymbol{\theta}) |\tau - \mathrm{I}(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}) < r(\boldsymbol{w}_i, \boldsymbol{\theta}))|$$
$$- r(\boldsymbol{w}_i, \boldsymbol{\theta})^2 |\tau - \mathrm{I}(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}) < r(\boldsymbol{w}_i, \boldsymbol{\theta}))| \Big]$$

Therefore, it can be obtained,

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} |l(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta})| \leq \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=1}^{n} (Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \right| |(1 - 2\tau)|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| 2 \sum_{i=1}^{N} r(\boldsymbol{w}_i, \boldsymbol{\theta})(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \mathrm{I}(|X_i - t| < \epsilon) \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| 2 \sum_{i=N+1}^{n} r(\boldsymbol{w}_i, \boldsymbol{\theta})(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \mathrm{I}(|X_i - t| < \epsilon) \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=1}^{n} r(\boldsymbol{w}_i, \boldsymbol{\theta})(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \mathrm{I}(|X_i - t| \geq \epsilon) \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=1}^{N} r(\boldsymbol{w}_i, \boldsymbol{\theta})^2 \mathrm{I}(|X_i - t| < \epsilon) \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=N+1}^{n} r(\boldsymbol{w}_i, \boldsymbol{\theta})^2 \mathrm{I}(|X_i - t| < \epsilon) \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=1}^{n} r(\boldsymbol{w}_i, \boldsymbol{\theta})^2 \mathrm{I}(|X_i - t| \geq \epsilon) \right|$$

Let $M = \max\{r(\boldsymbol{w}_i, \boldsymbol{\theta}), i = 1 \ldots, N\}$. Combining the existing assumption $\lim_{n \to \infty} P(|X_i - t| < \epsilon) = 0$ and the expansion $\Phi\left(\frac{X_i - t}{h}\right) = I(X_i > t) + o_i(h)$, we have,

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} |l(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta})| \leq \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=1}^{n} (Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \right| |(1 - 2\tau)|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| 2 \sum_{i=1}^{N} M \cdot (Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| 2 \sum_{i=N+1}^{n} r(\boldsymbol{w}_i, \boldsymbol{\theta})(Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \cdot \xi \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| 2 \sum_{i=1}^{n} (Y_i - v(\boldsymbol{w}_i, \boldsymbol{\theta}))^2 \beta_2(X_i - t) \cdot o_i(h) \cdot 1 \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=1}^{N} M^2 \cdot 1 \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=N+1}^{n} r(\boldsymbol{w}_i, \boldsymbol{\theta})^2 \cdot 1 \cdot \eta \right|$$

$$+ \sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \frac{1}{n} \left| \sum_{i=1}^{n} [\beta_2(X_i - t) \cdot o_i(h)]^2 \cdot 1 \right|$$

Obviously, the first term on the right side of the above equation tends to 0 when $n \to \infty$. For bounded N, the second, third and fifth term also tend to 0 when $n \to \infty$. ξ can be any small value, and by definition, $o_i(h)$ is also a small value, so we have,

$$\sup_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} |l(\boldsymbol{\theta}) - \ell(\boldsymbol{\theta})| \xrightarrow{p} 0$$

The consistency of the parameter estimators in the objective function $l_\tau(\boldsymbol{\theta})$ has been given by Newey and Powell [18], so the parameter estimators in the objective function $\ell_\tau(\boldsymbol{\theta})$ are also consistent. Thus, we complete the proof of Lemma A1. □

Lemma A2. *Define,*

$$u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0) = \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))q(\boldsymbol{w}_i, \boldsymbol{\theta}) - \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)$$

Under the conditions A1–A5, for any positive sequence d_n converging to 0 as n tends to infinity, we have,

$$\sup_{\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \le d_n} n^{-\frac{1}{2}} \| \sum_{i=1}^n \{u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0) - E[u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0)]\} \| = o_p(1)$$

Proof of Lemma A2. The above formula can be proved by Lemma 4.6 in He and Shao [23]. To prove Lemma A2, it is only necessary to state that the conditions (B1), (B3) and (B5′) of Lemma 4.6 in He and Shao [23] are satisfied.

For condition (B1), the measurability is clearly obvious.

For condition (B3), we have,

$$\| u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0(\tau)) \| \le \| \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))\{q(\boldsymbol{w}_i, \boldsymbol{\theta}) - q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)\} \|$$
$$+ \| \{\psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta})) - \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))\}q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$
$$\triangleq \| \mathrm{I}_{1i} \| + \| \mathrm{I}_{2i} \|$$

For I_{1i}, applying the mean value theorem and Lemma A1, we have,

$$\| \mathrm{I}_{1i} \| = \| \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))\{q(\boldsymbol{w}_i, \boldsymbol{\theta}) - q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)\} \|$$
$$= \| \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))q'(\boldsymbol{w}_i, \boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) \| = o_p(1)$$

Hence, $E\left(\| \mathrm{I}_{1i} \|^2 \mid \boldsymbol{w}_i\right) = o_p(1)$.

For I_{2i},

$$\| \mathrm{I}_{2i} \| = \| \{\psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta})) - \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))\}q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$
$$= \| [|\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))$$
$$- |\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))]q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$
$$\le \| [|\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))$$
$$- |\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))]q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$
$$\le \| |\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))|\{g(\boldsymbol{w}_i, \boldsymbol{\theta}) - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)\} \|$$
$$+ \| [|\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))| - |\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}))|](Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$
$$\triangleq \| \mathrm{II}_{1i} \| + \| \mathrm{II}_{2i} \|$$

For II_{1i}, based on Lemma A1, we have,

$$\| \mathrm{II}_{1i} \| = \| |\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))|\{g(\boldsymbol{w}_i, \boldsymbol{\theta}) - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)\} \| \le d_n \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$

Hence, $E\left(\| \mathrm{II}_{1i} \|^2 \mid \boldsymbol{w}_i\right) = o_p(1)$.

For II_{2i},

$$\| \mathrm{II}_{2i} \| = \| [|\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))| - |\tau - \mathrm{I}(Y_i \le g(\boldsymbol{w}_i, \boldsymbol{\theta}))|](Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$
$$\le \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)) \| \cdot |2\tau - 1| \cdot \mathrm{I}\{g_1(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0) \le Y_i \le g_2(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0)\}$$

where,

$$g_1(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0) = \min\{g(\boldsymbol{w}_i, \boldsymbol{\theta}), g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)\}$$
$$g_2(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0) = \max\{g(\boldsymbol{w}_i, \boldsymbol{\theta}), g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)\}$$

Obviously, $g_1(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0) \le g_2(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0)$. Depending on,

$$\| g(\boldsymbol{w}_i, \boldsymbol{\theta}) - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \| = \| q^\top(\boldsymbol{w}_i, \boldsymbol{\theta}_0)(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + R(n) \|$$
$$\leq \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \| \| \boldsymbol{\theta} - \boldsymbol{\theta}_0 \|$$
$$\leq d_n \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$

and the mean value theorem, we have,

$$E(\| \mathrm{II}_{2i} \|^2 \boldsymbol{w}_i) \leq (2\tau - 1)^2 \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|^2 \| Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|^2$$
$$* E[\mathrm{I}\{g_1(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0) \leq Y_i \leq g_2(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0)\}]$$
$$\leq (2\tau - 1)^2 \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|^2 \| Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|^2$$
$$* f_{\tau, \boldsymbol{w}_i}(\eta_i) \| g(\boldsymbol{w}_i, \boldsymbol{\theta}) - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|$$
$$\leq d_n f_{\tau, \boldsymbol{w}_i}(\eta_i)(2\tau - 1)^2 \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|^3 \| Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|^2$$

where η_i is a value between $g_1(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0)$ and $g_2(\boldsymbol{w}_i, \boldsymbol{\theta}, \boldsymbol{\theta}_0)$. In summary, we have,

$$E\left(\| u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0(\tau)) \|^2 \Big| \boldsymbol{w}_i \right) \leq a_i^2 d$$

where,

$$a_i = \sqrt{d_n \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \| \left[1 + f_{\tau, \boldsymbol{w}_i} n^{1/2}(2\tau - 1)^2 \| q(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|^2 \| Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \|^2 \right]}$$

and $d = n^{-1/2}$. Therefore, $u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ satisfies (B3).

For condition (B5'), let $A_n = \sum_{i=1}^n a_i^2$, for arbitrary constant C, we have,

$$P\left\{ \max_{1 \leq i \leq n} \| u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \| \geq C A_n^{1/2} d_n^{1/2}(\log n)^{-2} \right\}$$
$$\leq \sum_{i=1}^n P\left\{ \| u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \| \geq C A_n^{1/2} d_n^{1/2}(\log n)^{-2} \right\}$$
$$\leq \sum_{i=1}^n \frac{E(\| u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \|^2)}{C^2 A_n d_n(\log n)^{-4}} \leq \frac{n^{-1/2} A_n}{C^2 A_n d_n(\log n)^{-4}} = \frac{(\log n)^4}{C^2 d_n \sqrt{n}} = o_p(1)$$

therefore,

$$\max_{1 \leq i \leq n} \| u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \| = O_p\left(A_n^{1/2} d_n^{1/2} \left(\log n \right)^{-2} \right)$$

So $u_i(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ satisfies (B5'). Therefore, we complete the proof of Lemma A2. □

Proof of Theorem 1. By Lemma A1 and Lemma A2, we have,

$$n^{-\frac{1}{2}} \| \sum_{i=1}^n \psi_\tau \big(Y_i - g(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) \big) q(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) - \sum_{i=1}^n \psi_\tau \big(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0) \big) q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)$$
$$- \sum_{i=1}^n E\big[\psi_\tau \big(Y_i - g(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) \big) q(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) \big] \| = o_p(1)$$

By applying Taylor expansion for $\sum_{i=1}^n E\big[\psi_\tau \big(Y_i - g(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) \big) q(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) \big]$ around $\boldsymbol{\theta}_0$, we have,

$$E\left[\sum_{i=1}^n \psi_\tau \big(Y_i - g(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) \big) q(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) \right]$$
$$= \frac{\partial E[\sum_{i=1}^n \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))q(\boldsymbol{w}_i, \boldsymbol{\theta})]}{\partial \boldsymbol{\theta}} \bigg|_{\boldsymbol{\theta} = \boldsymbol{\theta}_0} \big[\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \big] + R_n$$
$$\triangleq n D_{n,\tau} \big[\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \big] + R_n$$

where $R_n = o_p(\sqrt{n})$ and,

$$D_h(\boldsymbol{\theta}_0) = n^{-1} \frac{\partial E[\sum_{i=1}^{n} \psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))q(\boldsymbol{w}_i, \boldsymbol{\theta})]}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

$$= n^{-1}\sum_{i=1}^{n} \frac{\partial E[\psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))q(\boldsymbol{w}_i, \boldsymbol{\theta})]}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

$$= n^{-1}\sum_{i=1}^{n} \frac{\partial E[|\tau - \mathrm{I}(Y_i < g(\boldsymbol{w}_i, \boldsymbol{\theta}))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))q(\boldsymbol{w}_i, \boldsymbol{\theta})]}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

$$= n^{-1}\sum_{i=1}^{n} \frac{\partial q(\boldsymbol{w}_i, \boldsymbol{\theta}) E[|\tau - \mathrm{I}(Y_i < g(\boldsymbol{w}_i, \boldsymbol{\theta}))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))]}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

$$= n^{-1}\sum_{i=1}^{n}\left[\frac{\partial q(\boldsymbol{w}_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} E[|\tau - \mathrm{I}(Y_i < g(\boldsymbol{w}_i, \boldsymbol{\theta}))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))]\right.$$

$$\left. + q(\boldsymbol{w}_i, \boldsymbol{\theta}) \frac{\partial E[|\tau - \mathrm{I}(Y_i < g(\boldsymbol{w}_i, \boldsymbol{\theta}))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))]}{\partial \boldsymbol{\theta}}\right]\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

$$= n^{-1}\sum_{i=1}^{n} q(\boldsymbol{w}_i, \boldsymbol{\theta}) \frac{\partial E[|\tau - \mathrm{I}(Y_i < g(\boldsymbol{w}_i, \boldsymbol{\theta}))|(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}))]}{\partial \boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$$

$$= -n^{-1}\sum_{i=1}^{n} q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)q^\top(\boldsymbol{w}_i, \boldsymbol{\theta}_0)[(1 - 2\tau)F_i(g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)) + \tau]$$

In addition, by $n^{-1/2}\sum_{i=1}^{n} \psi_\tau(Y_i - g(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n))q(\boldsymbol{w}_i, \hat{\boldsymbol{\theta}}_n) = o_p(1)$, we have,

$$-n^{-\frac{1}{2}}\sum_{i=1}^{n}\{\psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)\} - n^{-\frac{1}{2}}nD_h(\boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) - R_n = o_p(1)$$

All in all, we have,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = -n^{-1/2}D_h(\boldsymbol{\theta}_0)\sum_{i=1}^{n}\{\psi_\tau(Y_i - g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)\} + o_p(\sqrt{n})$$

which is equivalent to,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \to \mathcal{N}\left(0, [(1 - 2\tau)F_i(g(\boldsymbol{w}_i, \boldsymbol{\theta}_0)) + \tau]D_h(\boldsymbol{\theta}_0)^{-1}C_h(\boldsymbol{\theta}_0)D_h(\boldsymbol{\theta}_0)^{-\top}\right)$$

where $C_h(\boldsymbol{\theta}_0) = \frac{[(1-2\tau)F_i(g(\boldsymbol{w}_i, \boldsymbol{\theta}_0))+\tau]}{n}\sum_{i=1}^{n} q(\boldsymbol{w}_i, \boldsymbol{\theta}_0)q^\top(\boldsymbol{w}_i, \boldsymbol{\theta}_0)$. So far, the proof of the theorem is completed. $\square$

Appendix B. Tables of Results

Table A1. The simulation results of the independent identically distribution model with $\tilde{e} \sim N(0, 1)$ in simulation 1.

τ		Seg					Proposed				
		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
0.1	Mean	0.446	0.030	−0.031	−0.002	1.373	0.017	0.033	−0.033	0.001	1.371
	SSE	0.476	0.143	0.176	0.180	1.394	0.212	0.174	0.243	0.241	1.396
	Min	−0.092	−0.353	−0.460	−0.401	0.558	−0.714	−0.443	−0.649	−0.698	0.271
	IQ1	0.344	−0.057	−0.130	−0.132	1.266	−0.117	−0.090	−0.196	−0.171	1.252

Table A1. *Cont.*

τ				Seg					Proposed		
		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
	Median	0.430	0.027	−0.036	−0.002	1.370	0.001	0.038	−0.013	0.033	1.374
	IQ3	0.569	0.114	0.076	0.124	1.499	0.151	0.132	0.134	0.158	1.500
	Max	0.873	0.484	0.469	0.544	1.983	0.605	0.535	0.618	0.703	2.104
0.3	Mean	0.200	0.005	−0.007	0.013	1.198	0.005	0.001	0.006	0.010	1.196
	SSE	0.240	0.121	0.150	0.147	1.214	0.159	0.139	0.194	0.192	1.211
	Min	−0.130	−0.327	−0.442	−0.400	0.613	−0.389	−0.292	−0.509	−0.579	0.670
	IQ1	0.111	−0.081	−0.098	−0.078	1.092	−0.105	−0.109	−0.148	−0.118	1.089
	Median	0.189	0.002	−0.004	0.019	1.203	−0.007	−0.004	0.022	0.007	1.203
	IQ3	0.276	0.088	0.086	0.105	1.306	0.098	0.104	0.146	0.157	1.305
	Max	0.677	0.329	0.408	0.523	1.851	0.486	0.340	0.449	0.501	1.851
0.5	Mean	0.015	0.007	−0.014	−0.010	0.995	0.010	0.015	−0.020	−0.016	0.990
	SSE	0.126	0.107	0.133	0.148	1.012	0.162	0.131	0.183	0.178	1.007
	Min	−0.316	−0.347	−0.391	−0.449	0.500	−0.466	−0.450	−0.625	−0.609	0.477
	IQ1	−0.066	−0.064	−0.108	−0.099	0.881	−0.091	−0.077	−0.129	−0.123	0.878
	Median	0.025	0.003	−0.013	−0.015	0.995	0.007	0.023	−0.029	−0.018	0.988
	IQ3	0.098	0.081	0.077	0.096	1.110	0.121	0.102	0.099	0.091	1.107
	Max	0.364	0.268	0.516	0.385	1.589	0.417	0.343	0.711	0.442	1.589
0.7	Mean	−0.153	0.012	−0.009	−0.014	0.771	0.043	0.014	−0.011	−0.028	0.770
	SSE	0.215	0.106	0.137	0.146	0.792	0.175	0.114	0.166	0.186	0.790
	Min	−0.566	−0.275	−0.377	−0.468	0.240	−0.497	−0.264	−0.397	−0.589	0.238
	IQ1	−0.257	−0.063	−0.098	−0.113	0.654	−0.061	−0.065	−0.130	−0.155	0.662
	Median	−0.160	0.006	−0.004	−0.010	0.768	0.047	0.011	0.000	−0.034	0.764
	IQ3	−0.052	0.067	0.092	0.069	0.890	0.148	0.081	0.107	0.086	0.881
	Max	0.200	0.293	0.351	0.405	1.330	0.507	0.348	0.355	0.531	1.330
0.9	Mean	−0.448	−0.017	−0.011	−0.010	0.653	−0.016	−0.014	−0.019	−0.011	0.655
	SSE	0.471	0.129	0.172	0.169	0.685	0.206	0.153	0.237	0.255	0.686
	Min	−0.917	−0.389	−0.492	−0.404	0.020	−0.617	−0.477	−0.579	−0.638	0.021
	IQ1	−0.540	−0.084	−0.133	−0.114	0.524	−0.144	−0.103	−0.184	−0.169	0.536
	Median	−0.444	−0.015	−0.016	−0.011	0.645	−0.007	−0.025	−0.015	−0.026	0.645
	IQ3	−0.335	0.055	0.090	0.110	0.770	0.103	0.081	0.150	0.176	0.772
	Max	−0.057	0.406	0.557	0.389	1.319	0.627	0.478	0.694	0.659	1.343

Table A2. The simulation results of the independent identically distribution model with $\widetilde{e} \sim t_3$ in simulation 1.

τ				Seg					Proposed		
		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
0.1	Mean	0.364	0.063	−0.130	0.062	1.394	0.012	0.062	−0.095	0.073	1.361
	SSE	0.582	0.369	0.548	0.469	1.513	0.406	0.384	0.503	0.439	1.491
	Min	−1.629	−0.824	−3.523	−2.302	−0.464	−1.385	−0.733	−2.451	−1.499	−0.435
	IQ1	0.104	−0.151	−0.405	−0.236	1.114	−0.243	−0.202	−0.389	−0.217	1.051
	Median	0.403	0.023	−0.075	0.077	1.375	0.030	0.031	−0.021	0.103	1.353
	IQ3	0.654	0.272	0.225	0.338	1.708	0.254	0.254	0.228	0.359	1.681
	Max	1.536	1.625	1.035	1.486	3.562	1.197	1.992	1.139	1.280	3.662
0.3	Mean	0.156	0.038	−0.038	−0.024	1.168	0.025	0.027	−0.011	−0.026	1.168
	SSE	0.303	0.217	0.287	0.248	1.217	0.232	0.203	0.236	0.214	1.220
	Min	−0.459	−0.402	−1.092	−0.788	−0.564	−0.538	−0.417	−0.674	−0.543	−0.565
	IQ1	0.005	−0.106	−0.201	−0.196	0.957	−0.115	−0.124	−0.174	−0.179	0.971
	Median	0.146	0.010	−0.041	−0.010	1.182	0.029	0.016	−0.019	−0.026	1.167
	IQ3	0.309	0.179	0.144	0.160	1.386	0.151	0.174	0.146	0.109	1.378
	Max	1.692	1.162	0.641	0.512	2.307	1.342	0.812	0.650	0.626	2.307
0.5	Mean	0.020	0.009	−0.047	−0.006	1.026	0.009	−0.004	−0.018	0.000	1.030
	SSE	0.251	0.202	0.252	0.251	1.097	0.220	0.196	0.221	0.204	1.104

Table A2. *Cont.*

τ		**Seg**					**Proposed**				
		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
0.1	Mean	0.364	0.063	−0.130	0.062	1.394	0.012	0.062	−0.095	0.073	1.361
	SSE	0.582	0.369	0.548	0.469	1.513	0.406	0.384	0.503	0.439	1.491
	Min	−1.629	−0.824	−3.523	−2.302	−0.464	−1.385	−0.733	−2.451	−1.499	−0.435
	IQ1	0.104	−0.151	−0.405	−0.236	1.114	−0.243	−0.202	−0.389	−0.217	1.051
	Median	0.403	0.023	−0.075	0.077	1.375	0.030	0.031	−0.021	0.103	1.353
	IQ3	0.654	0.272	0.225	0.338	1.708	0.254	0.254	0.228	0.359	1.681
	Max	1.536	1.625	1.035	1.486	3.562	1.197	1.992	1.139	1.280	3.662
0.3	Mean	0.156	0.038	−0.038	−0.024	1.168	0.025	0.027	−0.011	−0.026	1.168
	SSE	0.303	0.217	0.287	0.248	1.217	0.232	0.203	0.236	0.214	1.220
	Min	−0.459	−0.402	−1.092	−0.788	−0.564	−0.538	−0.417	−0.674	−0.543	−0.565
	IQ1	0.005	−0.106	−0.201	−0.196	0.957	−0.115	−0.124	−0.174	−0.179	0.971
	Median	0.146	0.010	−0.041	−0.010	1.182	0.029	0.016	−0.019	−0.026	1.167
	IQ3	0.309	0.179	0.144	0.160	1.386	0.151	0.174	0.146	0.109	1.378
	Max	1.692	1.162	0.641	0.512	2.307	1.342	0.812	0.650	0.626	2.307
0.5	Mean	0.020	0.009	−0.047	−0.006	1.026	0.009	−0.004	−0.018	0.000	1.030
	SSE	0.251	0.202	0.252	0.251	1.097	0.220	0.196	0.221	0.204	1.104
	Min	−0.793	−0.704	−0.913	−0.914	−0.545	−0.601	−0.882	−0.896	−0.430	−0.684
	IQ1	−0.117	−0.105	−0.223	−0.154	0.842	−0.128	−0.119	−0.137	−0.142	0.845
	Median	0.009	−0.002	−0.062	−0.015	1.009	0.000	−0.006	−0.023	0.005	1.014
	IQ3	0.140	0.117	0.157	0.155	1.150	0.118	0.103	0.122	0.146	1.162
	Max	1.289	0.744	0.521	0.697	3.143	1.166	0.678	0.569	0.531	3.170
0.7	Mean	−0.062	0.064	−0.124	0.012	0.763	0.063	0.052	−0.062	0.011	0.751
	SSE	0.325	0.276	0.634	0.276	0.867	0.281	0.229	0.265	0.233	0.825
	Min	−0.745	−0.801	−7.756	−0.748	−0.825	−0.556	−0.415	−1.368	−0.517	−0.902
	IQ1	−0.249	−0.080	−0.239	−0.181	0.560	−0.096	−0.083	−0.207	−0.149	0.578
	Median	−0.095	0.058	−0.073	0.009	0.765	0.040	0.032	−0.056	0.010	0.761
	IQ3	0.079	0.175	0.107	0.181	0.981	0.200	0.169	0.107	0.164	0.982
	Max	2.582	2.013	0.599	0.936	3.613	2.079	1.833	0.597	0.580	1.492
0.9	Mean	−0.534	−0.072	−0.045	0.000	0.566	0.153	0.097	−0.231	0.003	0.606
	SSE	4.936	2.794	2.942	0.480	0.890	0.556	0.430	0.794	0.461	0.897
	Min	−68.813	−38.697	−8.077	−1.187	−2.735	−0.730	−1.108	−7.872	−1.376	−1.378
	IQ1	−0.583	−0.177	−0.410	−0.305	0.236	−0.183	−0.153	−0.440	−0.273	0.249
	Median	−0.326	0.054	−0.171	0.001	0.599	0.062	0.046	−0.161	−0.019	0.573
	IQ3	−0.010	0.337	0.099	0.260	0.945	0.434	0.316	0.093	0.273	0.916
	Max	4.566	3.347	39.707	3.021	3.283	3.237	2.345	0.942	1.315	3.283

Table A3. The simulation results of the independent identically distribution model with $\widetilde{e} \sim 0.9N(0,1) + 0.1t_3$ in simulation 1.

τ		**Seg**					**Proposed**				
		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
0.1	Mean	0.542	0.005	−0.020	−0.001	1.412	0.153	0.001	−0.015	0.011	1.412
	SSE	0.561	0.119	0.158	0.156	1.428	0.248	0.148	0.229	0.222	1.427
	Min	0.147	−0.279	−0.475	−0.505	0.686	−0.293	−0.365	−0.753	−0.713	0.683
	IQ1	0.456	−0.073	−0.108	−0.093	1.306	0.004	−0.098	−0.139	−0.130	1.308
	Median	0.537	−0.005	−0.026	0.011	1.413	0.162	−0.018	−0.023	0.031	1.408
	IQ3	0.624	0.073	0.067	0.099	1.530	0.278	0.082	0.141	0.154	1.528
	Max	0.920	0.547	0.490	0.420	1.873	0.646	0.651	0.772	0.580	1.873
0.3	Mean	0.221	0.001	−0.013	−0.003	1.214	0.058	−0.007	−0.008	−0.007	1.217
	SSE	0.255	0.102	0.136	0.152	1.227	0.166	0.112	0.162	0.186	1.230
	Min	−0.259	−0.273	−0.327	−0.344	0.763	−0.558	−0.291	−0.443	−0.523	0.763
	IQ1	0.136	−0.060	−0.104	−0.117	1.132	−0.026	−0.084	−0.113	−0.133	1.132
	Median	0.227	0.001	−0.008	−0.004	1.213	0.073	−0.011	−0.007	−0.013	1.215
	IQ3	0.298	0.059	0.076	0.105	1.308	0.147	0.065	0.096	0.098	1.309
	Max	0.584	0.351	0.380	0.389	1.595	0.578	0.301	0.437	0.630	1.596

Table A3. *Cont.*

				Seg					Proposed		
τ		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
0.5	Mean	0.004	−0.008	−0.004	−0.013	1.016	−0.007	−0.004	−0.009	−0.006	1.017
	SSE	0.110	0.096	0.131	0.132	1.029	0.135	0.106	0.163	0.154	1.030
	Min	−0.254	−0.348	−0.367	−0.469	0.650	−0.365	−0.385	−0.444	−0.465	0.689
	IQ1	−0.062	−0.071	−0.097	−0.098	0.917	−0.093	−0.080	−0.130	−0.108	0.919
	Median	0.002	−0.009	0.000	−0.019	1.006	−0.002	−0.002	0.001	−0.002	1.007
	IQ3	0.082	0.063	0.091	0.072	1.113	0.079	0.076	0.095	0.082	1.110
	Max	0.364	0.255	0.395	0.343	1.548	0.414	0.234	0.481	0.456	1.548
0.7	Mean	−0.194	0.003	−0.016	−0.018	0.799	−0.020	0.003	−0.013	−0.016	0.797
	SSE	0.234	0.110	0.147	0.140	0.818	0.145	0.122	0.183	0.174	0.815
	Min	−0.519	−0.265	−0.430	−0.428	0.215	−0.378	−0.274	−0.589	−0.429	0.216
	IQ1	−0.275	−0.082	−0.115	−0.121	0.693	−0.120	−0.082	−0.124	−0.136	0.698
	Median	−0.195	0.007	−0.001	−0.007	0.796	−0.027	0.009	−0.018	−0.012	0.797
	IQ3	−0.118	0.066	0.089	0.070	0.900	0.080	0.069	0.114	0.103	0.894
	Max	0.595	0.456	0.307	0.313	1.297	0.406	0.361	0.405	0.395	1.297
0.9	Mean	−0.547	−0.005	−0.015	0.009	0.626	−0.150	−0.008	−0.025	0.002	0.630
	SSE	0.569	0.120	0.174	0.182	0.648	0.243	0.147	0.236	0.242	0.652
	Min	−0.899	−0.326	−0.433	−0.485	0.226	−0.723	−0.465	−0.647	−0.608	0.236
	IQ1	−0.658	−0.088	−0.130	−0.139	0.532	−0.288	−0.116	−0.197	−0.173	0.526
	Median	−0.553	−0.001	−0.004	0.016	0.636	−0.146	−0.003	−0.030	−0.020	0.636
	IQ3	−0.445	0.079	0.114	0.126	0.741	−0.014	0.078	0.141	0.195	0.740
	Max	−0.131	0.375	0.412	0.492	1.030	0.362	0.372	0.550	0.603	1.039

Table A4. The simulation results of the heteroscedasticity model with $\widetilde{e} \sim N(0,1)$ in simulation 1.

				Seg					Proposed		
τ		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
0.1	Mean	0.400	0.126	−0.141	0.021	1.504	−0.023	0.003	−0.079	0.035	1.462
	SSE	0.467	0.210	0.784	0.293	1.585	0.291	0.167	0.522	0.310	1.521
	Min	−0.200	−0.593	−9.272	−0.791	0.590	−1.034	−0.549	−4.145	−0.922	0.590
	IQ1	0.259	0.045	−0.264	−0.174	1.267	−0.185	−0.106	−0.308	−0.159	1.247
	Median	0.408	0.137	−0.057	0.008	1.411	−0.014	−0.005	−0.022	0.046	1.402
	IQ3	0.572	0.223	0.142	0.250	1.614	0.185	0.111	0.242	0.226	1.592
	Max	1.036	0.496	0.670	0.937	4.059	0.902	0.584	0.886	1.059	3.959
0.3	Mean	0.186	0.063	−0.036	−0.011	1.240	−0.003	−0.002	−0.043	−0.021	1.257
	SSE	0.257	0.133	0.224	0.246	1.274	0.189	0.124	0.304	0.231	1.299
	Min	−0.483	−0.305	−0.879	−0.672	0.609	−0.601	−0.446	−1.316	−0.656	0.615
	IQ1	0.073	−0.010	−0.168	−0.183	1.077	−0.119	−0.088	−0.224	−0.175	1.072
	Median	0.188	0.069	−0.060	0.002	1.201	0.000	0.007	−0.033	0.000	1.205
	IQ3	0.315	0.138	0.108	0.139	1.400	0.138	0.079	0.158	0.131	1.404
	Max	0.576	0.483	0.662	0.734	2.460	0.471	0.527	0.818	0.606	3.157
0.5	Mean	−0.003	0.017	−0.004	0.038	0.966	0.019	0.018	−0.018	0.005	0.970
	SSE	0.168	0.120	0.222	0.243	1.011	0.184	0.137	0.292	0.210	1.015
	Min	−0.559	−0.289	−0.580	−0.555	−0.073	−0.461	−0.304	−0.837	−0.525	−0.076
	IQ1	−0.103	−0.067	−0.156	−0.108	0.777	−0.105	−0.065	−0.199	−0.134	0.809
	Median	−0.011	0.017	−0.010	0.023	0.959	0.015	0.006	−0.005	0.009	0.958
	IQ3	0.091	0.102	0.155	0.193	1.149	0.137	0.103	0.176	0.146	1.157
	Max	0.473	0.378	0.555	0.663	1.833	0.542	0.448	0.581	0.553	1.838
0.7	Mean	−0.180	−0.048	−0.024	0.004	0.795	0.000	0.008	−0.019	0.009	0.799
	SSE	0.256	0.130	0.242	0.236	0.839	0.194	0.131	0.293	0.234	0.841
	Min	−0.768	−0.411	−0.802	−0.638	0.122	−0.680	−0.361	−1.018	−0.753	0.059
	IQ1	−0.300	−0.131	−0.143	−0.171	0.642	−0.133	−0.085	−0.206	−0.152	0.644

Table A4. *Cont.*

τ		Seg					Proposed				
		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
	Median	−0.198	−0.050	−0.025	0.010	0.803	−0.006	0.010	−0.030	0.029	0.801
	IQ3	−0.058	0.033	0.132	0.179	0.929	0.111	0.087	0.178	0.176	0.929
	Max	0.437	0.407	0.600	0.579	2.037	0.634	0.479	0.684	0.578	2.037
0.9	Mean	−0.432	−0.152	−0.053	−0.017	0.676	−0.001	−0.023	−0.057	−0.034	0.668
	SSE	0.482	0.214	0.367	0.269	0.781	0.255	0.149	0.395	0.296	0.756
	Min	−0.998	−0.854	−3.220	−0.613	−0.230	−0.753	−0.450	−1.445	−0.693	−0.270
	IQ1	−0.578	−0.252	−0.214	−0.178	0.454	−0.173	−0.146	−0.250	−0.227	0.458
	Median	−0.429	−0.134	−0.009	−0.011	0.654	0.002	−0.012	−0.023	−0.029	0.637
	IQ3	−0.297	−0.043	0.145	0.165	0.828	0.176	0.093	0.182	0.166	0.808
	Max	0.222	0.158	0.698	0.723	2.916	0.674	0.282	1.098	0.763	2.254

Table A5. The simulation results of the heteroscedasticity model with $\tilde{e} \sim t_3$ in simulation 1.

τ		Seg					Proposed				
		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
0.1	Mean	0.490	0.270	−1.561	0.044	1.445	1.737	0.958	−2.496	0.036	1.401
	SSE	2.003	1.349	11.857	0.757	1.734	24.604	12.851	20.379	0.506	1.678
	Min	−3.847	−1.192	−98.359	−1.662	−1.459	−2.236	−0.748	−91.346	−1.081	−2.009
	IQ1	0.090	0.021	−0.593	−0.447	0.883	−0.301	−0.168	−0.652	−0.293	0.860
	Median	0.413	0.200	−0.151	−0.043	1.355	0.030	0.046	−0.127	−0.020	1.323
	IQ3	0.776	0.357	0.240	0.464	1.782	0.333	0.244	0.289	0.373	1.747
	Max	26.441	18.120	4.681	4.545	4.657	347.012	181.233	1.389	2.174	4.730
0.3	Mean	0.093	0.020	−0.228	−0.016	1.302	0.007	0.001	−0.170	−0.027	1.236
	SSE	0.395	0.231	1.462	0.461	1.451	0.282	0.186	1.345	0.273	1.369
	Min	−3.043	−0.744	−17.518	−1.200	−0.091	−0.794	−0.598	−17.991	−0.733	−0.690
	IQ1	−0.070	−0.108	−0.357	−0.318	0.951	−0.170	−0.098	−0.263	−0.187	0.946
	Median	0.109	0.035	−0.052	−0.020	1.205	0.018	−0.001	−0.062	−0.039	1.185
	IQ3	0.318	0.158	0.235	0.277	1.512	0.181	0.121	0.185	0.140	1.440
	Max	0.832	0.709	0.979	2.286	4.253	1.389	0.721	0.863	0.978	4.139
0.5	Mean	−0.003	−0.007	−0.333	0.026	1.172	−0.012	−0.020	−0.205	0.019	1.127
	SSE	0.546	0.367	1.143	0.387	1.377	0.285	0.204	0.793	0.232	1.300
	Min	−0.721	−0.786	−7.182	−1.134	−0.908	−0.731	−0.771	−5.516	−0.601	−0.588
	IQ1	−0.280	−0.167	−0.444	−0.204	0.797	−0.167	−0.110	−0.301	−0.121	0.729
	Median	−0.048	−0.002	−0.108	0.066	1.064	−0.032	−0.034	−0.049	0.036	1.072
	IQ3	0.168	0.120	0.162	0.276	1.404	0.128	0.079	0.145	0.171	1.341
	Max	6.292	3.941	1.008	0.908	3.694	1.695	1.108	0.984	0.647	3.610
0.7	Mean	−0.057	−0.035	−0.296	−0.022	0.907	0.065	0.010	−0.339	−0.034	0.908
	SSE	0.354	0.288	1.261	0.393	1.175	0.324	0.238	1.749	0.266	1.172
	Min	−0.816	−1.023	−12.200	−0.979	−0.398	−0.727	−0.788	−19.732	−0.724	−0.895
	IQ1	−0.289	−0.208	−0.403	−0.308	0.452	−0.133	−0.123	−0.338	−0.197	0.456
	Median	−0.103	−0.034	−0.086	−0.041	0.782	0.030	0.019	−0.110	−0.045	0.784
	IQ3	0.165	0.124	0.161	0.249	1.175	0.237	0.136	0.104	0.150	1.198
	Max	1.471	0.955	0.972	1.330	3.780	1.912	1.177	1.080	0.730	3.861
0.9	Mean	−0.126	−0.059	0.189	−0.054	0.915	0.092	0.001	1.195	0.000	0.800
	SSE	1.096	0.698	6.296	0.842	1.439	0.496	0.331	19.240	0.496	1.235
	Min	−1.327	−1.387	−20.019	−5.030	−1.913	−0.949	−0.950	−6.723	−1.326	−1.240
	IQ1	−0.651	−0.388	−0.996	−0.510	0.271	−0.226	−0.194	−0.706	−0.334	0.269
	Median	−0.315	−0.169	−0.302	−0.058	0.733	0.069	−0.001	−0.224	−0.019	0.637
	IQ3	0.092	0.113	0.213	0.436	1.365	0.360	0.150	0.252	0.289	1.306
	Max	7.680	4.569	42.983	3.122	3.799	3.011	1.803	266.968	1.463	3.931

Table A6. The simulation results of the heteroscedasticity model with $\widetilde{e} \sim 0.9N(0,1) + 0.1t_3$ in simulation 1.

τ				Seg					Proposed		
		β_0	β_1	β_2	γ	t	β_0	β_1	β_2	γ	t
0.1	Mean	0.522	0.166	−0.054	0.017	1.448	0.143	0.045	−0.053	0.005	1.453
	SSE	0.554	0.206	0.273	0.248	1.483	0.254	0.147	0.379	0.265	1.488
	Min	0.019	−0.197	−0.739	−0.705	0.626	−0.460	−0.340	−0.998	−0.740	0.625
	IQ1	0.406	0.087	−0.212	−0.153	1.257	0.015	−0.056	−0.297	−0.174	1.261
	Median	0.526	0.174	−0.050	0.019	1.455	0.143	0.047	−0.030	0.018	1.459
	IQ3	0.641	0.257	0.115	0.202	1.609	0.276	0.144	0.209	0.170	1.617
	Max	1.083	0.454	0.670	0.723	2.477	0.818	0.412	0.767	0.634	2.476
0.3	Mean	0.211	0.056	−0.041	0.003	1.253	0.040	0.009	−0.035	0.014	1.250
	SSE	0.269	0.125	0.213	0.208	1.290	0.180	0.107	0.248	0.190	1.284
	Min	−0.355	−0.286	−0.943	−0.614	0.509	−0.429	−0.300	−1.181	−0.464	0.477
	IQ1	0.111	−0.005	−0.195	−0.145	1.089	−0.087	−0.058	−0.174	−0.121	1.088
	Median	0.207	0.067	−0.030	0.001	1.221	0.044	0.002	−0.009	0.011	1.227
	IQ3	0.324	0.133	0.100	0.140	1.380	0.157	0.091	0.114	0.143	1.381
	Max	0.672	0.290	0.509	0.645	2.445	0.480	0.249	0.555	0.493	2.270
0.5	Mean	−0.009	−0.007	−0.024	0.014	1.028	0.003	0.000	−0.043	0.007	1.029
	SSE	0.157	0.096	0.201	0.211	1.053	0.172	0.103	0.263	0.194	1.053
	Min	−0.383	−0.208	−0.497	−0.481	0.520	−0.434	−0.290	−0.788	−0.568	0.501
	IQ1	−0.112	−0.073	−0.152	−0.122	0.887	−0.097	−0.063	−0.229	−0.135	0.896
	Median	−0.002	−0.007	−0.043	0.005	1.024	0.024	0.010	−0.058	0.018	1.029
	IQ3	0.086	0.060	0.112	0.155	1.155	0.112	0.075	0.132	0.129	1.155
	Max	0.387	0.230	0.485	0.535	1.861	0.467	0.272	0.619	0.531	1.836
0.7	Mean	−0.234	−0.073	−0.039	0.011	0.840	−0.062	−0.019	−0.057	0.005	0.843
	SSE	0.282	0.129	0.195	0.193	0.881	0.184	0.111	0.255	0.185	0.884
	Min	−0.640	−0.369	−0.561	−0.466	0.111	−0.501	−0.346	−0.896	−0.450	0.111
	IQ1	−0.338	−0.137	−0.144	−0.126	0.691	−0.180	−0.090	−0.235	−0.119	0.682
	Median	−0.244	−0.073	−0.039	0.018	0.823	−0.066	−0.016	−0.053	−0.001	0.831
	IQ3	−0.118	0.003	0.074	0.144	0.987	0.047	0.058	0.123	0.124	0.996
	Max	0.315	0.195	0.414	0.575	1.644	0.529	0.246	0.589	0.518	1.644
0.9	Mean	−0.525	−0.169	−0.014	0.012	0.593	−0.150	−0.058	−0.016	0.023	0.600
	SSE	0.560	0.222	0.250	0.282	0.674	0.268	0.165	0.356	0.275	0.682
	Min	−1.093	−0.678	−0.732	−0.631	−0.455	−0.612	−0.662	−1.109	−0.634	−0.455
	IQ1	−0.671	−0.248	−0.166	−0.165	0.409	−0.309	−0.164	−0.262	−0.174	0.420
	Median	−0.538	−0.155	−0.022	−0.001	0.587	−0.157	−0.066	−0.014	0.010	0.585
	IQ3	−0.411	−0.096	0.166	0.171	0.732	−0.025	0.033	0.240	0.196	0.728
	Max	0.150	0.321	0.677	0.881	2.257	0.554	0.452	0.949	1.067	2.400

Table A7. The simulation results of the symmetric scenario in simulation 2.

τ			Seg				Proposed		
		β_0	β_1	β_2	t	β_0	β_1	β_2	t
0.1	Mean	3.958	0.792	−0.004	4.893	3.761	0.747	−0.009	4.900
	SSE	3.972	0.798	0.125	4.910	3.777	0.754	0.169	4.917
	Min	3.390	0.590	−0.460	4.280	3.255	0.534	−0.651	4.280
	IQ1	3.864	0.743	−0.080	4.759	3.597	0.688	−0.110	4.761
	Median	3.962	0.795	0.002	4.873	3.773	0.741	0.003	4.883
	IQ3	4.069	0.830	0.078	5.011	3.906	0.810	0.108	5.020
	Max	4.436	1.068	0.279	5.493	4.342	1.016	0.416	5.493
0.3	Mean	4.217	0.839	−0.006	4.706	4.121	0.819	−0.002	4.708
	SSE	4.230	0.843	0.103	4.721	4.134	0.823	0.132	4.722
	Min	3.921	0.700	−0.250	4.328	3.691	0.655	−0.351	4.328
	IQ1	4.134	0.803	−0.090	4.602	4.013	0.777	−0.105	4.603

Table A7. *Cont.*

		Seg				Proposed			
τ		β_0	β_1	β_2	t	β_0	β_1	β_2	t
	Median	4.206	0.842	−0.002	4.693	4.124	0.817	−0.010	4.694
	IQ3	4.308	0.874	0.063	4.788	4.206	0.863	0.092	4.786
	Max	4.632	0.974	0.362	5.275	4.599	1.006	0.371	5.275
0.5	Mean	4.366	0.879	−0.001	4.494	4.358	0.880	0.001	4.495
	SSE	4.379	0.883	0.114	4.508	4.371	0.884	0.125	4.509
	Min	4.012	0.752	−0.334	4.050	3.830	0.740	−0.354	4.052
	IQ1	4.279	0.838	−0.080	4.382	4.242	0.831	−0.090	4.379
	Median	4.366	0.879	0.001	4.498	4.357	0.880	−0.005	4.500
	IQ3	4.438	0.920	0.080	4.589	4.463	0.926	0.093	4.590
	Max	4.755	1.026	0.268	4.935	4.767	1.072	0.381	4.935
0.7	Mean	4.531	0.913	−0.007	4.296	4.616	0.933	−0.014	4.303
	SSE	4.544	0.916	0.102	4.310	4.629	0.938	0.119	4.317
	Min	4.145	0.796	−0.290	3.856	4.211	0.776	−0.361	3.856
	IQ1	4.456	0.876	−0.078	4.201	4.526	0.897	−0.102	4.200
	Median	4.524	0.914	−0.009	4.292	4.604	0.932	−0.016	4.294
	IQ3	4.598	0.946	0.060	4.406	4.718	0.969	0.076	4.409
	Max	4.883	1.124	0.270	4.742	5.065	1.133	0.264	4.742
0.9	Mean	4.793	0.963	−0.018	4.114	5.003	1.004	−0.015	4.117
	SSE	4.808	0.968	0.126	4.129	5.020	1.009	0.170	4.131
	Min	4.387	0.789	−0.385	3.454	4.510	0.825	−0.505	3.528
	IQ1	4.685	0.917	−0.099	4.003	4.868	0.953	−0.125	4.009
	Median	4.809	0.958	−0.017	4.109	5.008	0.998	−0.020	4.114
	IQ3	4.882	1.002	0.052	4.234	5.133	1.055	0.084	4.237
	Max	5.192	1.185	0.339	4.659	5.545	1.221	0.437	4.659

Table A8. The simulation results of the asymmetric scenario in simulation 2.

		Seg				Proposed			
τ		β_0	β_1	β_2	t	β_0	β_1	β_2	t
0.1	Mean	4.307	1.009	−0.508	2.564	4.120	1.385	−0.601	2.441
	SSE	4.406	1.638	1.929	3.253	4.203	8.160	8.039	2.840
	Min	−5.072	−0.427	−11.460	0.018	−4.462	−0.404	−93.116	−0.017
	IQ1	4.107	0.533	−0.514	1.575	3.886	0.385	−0.428	1.612
	Median	4.381	0.875	−0.112	1.910	4.224	0.847	−0.027	1.968
	IQ3	4.644	1.251	0.195	2.426	4.505	1.218	0.376	2.625
	Max	5.606	11.926	7.420	9.206	5.110	114.223	1.325	8.829
0.3	Mean	4.762	0.961	−0.543	2.043	4.712	0.931	−0.053	1.985
	SSE	4.789	1.062	4.243	2.494	4.735	1.025	0.643	2.302
	Min	3.512	−0.317	−51.275	0.723	3.927	−0.273	−7.557	0.706
	IQ1	4.507	0.742	−0.270	1.482	4.517	0.716	−0.238	1.486
	Median	4.768	0.971	−0.031	1.680	4.700	0.928	0.021	1.694
	IQ3	4.919	1.207	0.172	1.965	4.908	1.175	0.230	1.979
	Max	6.137	2.557	0.907	9.284	5.827	2.274	0.919	9.203
0.5	Mean	4.970	1.035	−0.039	1.553	5.001	1.030	−0.034	1.554
	SSE	4.989	1.087	0.307	1.634	5.019	1.084	0.314	1.643
	Min	4.035	0.098	−1.069	0.704	4.342	−0.011	−1.202	0.720
	IQ1	4.821	0.804	−0.205	1.279	4.863	0.831	−0.203	1.277
	Median	4.964	1.025	−0.020	1.475	4.990	1.033	−0.015	1.480
	IQ3	5.125	1.233	0.169	1.724	5.144	1.239	0.147	1.724
	Max	5.870	2.095	0.769	4.584	5.947	2.150	0.850	5.509
0.7	Mean	5.240	1.054	−0.044	1.572	5.324	1.057	−0.017	1.518
	SSE	5.263	1.150	0.457	1.894	5.346	1.137	0.404	1.747
	Min	4.542	−0.217	−2.073	0.159	4.601	−0.186	−1.984	0.283

Table A8. *Cont.*

τ		Seg				Proposed			
		β_0	β_1	β_2	t	β_0	β_1	β_2	t
	IQ1	5.065	0.848	−0.244	1.152	5.149	0.876	−0.200	1.157
	Median	5.192	1.038	0.000	1.353	5.280	1.031	0.007	1.358
	IQ3	5.367	1.257	0.184	1.635	5.453	1.220	0.186	1.624
	Max	6.921	3.212	2.800	7.803	6.823	2.969	1.617	7.516
0.9	Mean	5.733	1.106	0.041	1.783	6.062	1.003	0.144	2.119
	SSE	5.781	1.412	1.375	2.394	6.118	1.197	1.138	2.851
	Min	4.005	−0.329	−7.884	−0.487	4.801	−0.267	−2.182	−0.054
	IQ1	5.343	0.674	−0.325	0.924	5.584	0.618	−0.272	1.013
	Median	5.661	1.015	0.034	1.246	5.898	1.012	0.030	1.370
	IQ3	5.956	1.384	0.346	1.967	6.333	1.369	0.386	2.305
	Max	7.902	9.237	11.567	8.088	8.096	3.326	11.517	8.352

Table A9. The annual average temperature in the Arctic region and time.

τ	Method		β_0	β_1	β_2	t
0.2	Grid	Estimate	13.029	−0.007	0.035	1975.425
		SE	0.000	0.000	0.001	0.000
	Seg	Estimate	12.922	−0.007	0.035	1975.493
		SE	0.014	0.001	0.002	1.476
	proposed	Estimate	12.922	−0.007	0.035	1975.493
		SE	0.000	0.000	0.001	0.000
0.3	Grid	Estimate	13.911	−0.007	0.035	1974.748
		SE	0.000	0.000	0.001	0.000
	Seg	Estimate	13.727	−0.007	0.035	1974.857
		SE	0.015	0.002	0.002	1.582
	proposed	Estimate	13.170	−0.007	0.035	1975.204
		SE	0.000	0.000	0.001	0.000
0.5	Grid	Estimate	13.384	−0.007	0.035	1974.748
		SE	0.000	0.000	0.002	0.000
	Seg	Estimate	13.560	−0.007	0.035	1974.643
		SE	0.014	0.002	0.002	1.587
	proposed	Estimate	13.560	−0.007	0.035	1974.643
		SE	0.000	0.000	0.002	0.000
0.7	Grid	Estimate	12.873	−0.007	0.034	1974.748
		SE	0.000	0.000	0.001	0.000
	Seg	Estimate	13.105	−0.007	0.034	1974.612
		SE	0.014	0.002	0.002	1.708
	proposed	Estimate	13.105	−0.007	0.034	1974.611
		SE	0.000	0.000	0.001	0.000
0.8	Grid	Estimate	12.458	−0.006	0.034	1974.748
		SE	0.000	0.000	0.001	0.000
	Seg	Estimate	12.754	−0.006	0.034	1974.576
		SE	0.015	0.002	0.003	1.781
	proposed	Estimate	12.754	−0.006	0.034	1974.576
		SE	0.000	0.000	0.001	0.000

Table A10. The annual average flow of the River Nile and time.

τ	Method		β_0	β_1	β_2	t
0.2	Grid	Estimate	1100.048	−8.002	9.053	42.657
		SE	52.296	2.266	2.965	8.115
	Seg	Estimate	1101.656	−8.175	9.035	40.468
		SE	15.497	2.713	2.957	7.930
	proposed	Estimate	1099.208	−7.942	9.027	42.995
		SE	46.481	1.658	2.975	0.000

Table A10. *Cont.*

τ	Method		β_0	β_1	β_2	t
0.3	Grid	Estimate	1133.474	−8.130	9.069	42.657
		SE	51.070	2.312	2.882	8.362
	Seg	Estimate	1149.104	−9.198	9.133	38.146
		SE	13.803	1.994	2.234	6.649
	proposed	Estimate	1132.675	−8.072	9.047	42.997
		SE	47.380	1.511	2.737	0.000
0.5	Grid	Estimate	1183.102	−8.079	8.874	43.483
		SE	43.630	2.161	2.437	8.974
	Seg	Estimate	1168.128	−7.176	8.461	48.676
		SE	14.199	1.834	2.118	7.315
	proposed	Estimate	1184.582	−8.174	8.925	43.001
		SE	43.908	1.219	1.999	0.000
0.7	Grid	Estimate	1213.501	−6.889	8.327	51.738
		SE	37.726	2.371	2.540	12.134
	Seg	Estimate	1214.131	−6.924	8.327	51.471
		SE	15.375	1.738	2.242	7.687
	proposed	Estimate	1214.131	−6.924	8.327	51.471
		SE	32.335	0.870	1.655	0.000
0.8	Grid	Estimate	1241.886	−6.711	8.274	53.390
		SE	39.351	1.489	2.073	8.418
	Seg	Estimate	1241.697	−6.701	8.276	53.474
		SE	17.326	1.849	2.751	9.226
	proposed	Estimate	1241.697	−6.701	8.276	53.474
		SE	37.032	1.066	2.070	0.000

References

1. Zeileis, A. Implementing a class of structural change tests: An econometric computing approach. *Comput. Stat. Data Anal.* **2006**, *50*, 2987–3008. [CrossRef]
2. Goesmann, J.; Ziggel, D. An innovative risk management methodology for trading equity indices based on change points. *J. Asset Manag.* **2018**, *19*, 99–109. [CrossRef]
3. Guo, Y.Q.; Wang, X.Y.; Xu, Q.; Liu, F.F.; Liu, Y.Q.; Xia, Y.Y. Change-Point Analysis of Eye Movement Characteristics for Female Drivers in Anxiety. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1236. [CrossRef] [PubMed]
4. Mackintosh, M.; Aldridge, R.W.; Rossor, M.; Gonzalez-lzquierdo, A.; Whitaker, K.J.; Ford, E.; Direk, K.; Denaxas, S. Dementia recognition, diagnosis, and treatment in the UK, 1997–2007: A change-point analysis. *Lancet* **2019**, *394*, S70. [CrossRef]
5. Piegorsch, W.W.; Bailer, A.J. *Analyzing Environmental Data*; John Wiley & Sons: Hoboken, NJ, USA, 2005.
6. Xie, Y.; Wilson, A.M. Change point estimation of deciduous forest land surface phenology. *Remote Sens. Environ.* **2020**, *240*, 111698. [CrossRef]
7. Page, E.S. Continuous inspection schemes. *Biometrika* **1954**, *42*, 100–115. [CrossRef]
8. Cardoso-Silva, J.; Papadatos, G.; Papageorgiou, L.G.; Tsoka, S. Optimal Piecewise Linear Regression Algorithm for QSAR Modelling. *Mol. Inform.* **2019**, *38*, 1743–1868. [CrossRef] [PubMed]
9. Shi, S.; Li, Y.; Wan, C. Robust continuous piecewise linear regression model with multiple change points. *J. Supercomput.* **2020**, *76*, 3623–3645. [CrossRef]
10. Pakdaman, N.M.; Cooper, G.F. Binary classifier calibration using an ensemble of piecewise linear regression models. *Knowl. Inf. Syst.* **2018**, *54*, 151–170. [CrossRef]
11. Bodhlyera, O.; Zewotir, T.; Ramroop, S.; Chunilall, V. Analysis of the changes in chemical properties of dissolving pulp during the bleaching process using piecewise linear regression models. *Cell. Chem. Technol.* **2015**, *49*, 317–332.
12. Bhattacharya, P.K. Some aspects of change-point analysis. *Lect. Notes-Monogr. Ser.* **1994**, *23*, 28–56.
13. Chan, K.S. Testing for threshold autoregression. *Ann. Stat.* **1990**, *18*, 1886–1894. [CrossRef]
14. Liu, J.; Qian, L. Changepoint estimation in a segmented linear regression via empirical likelihood. *Commun. Stat.-Simul. Comput.* **2009**, *39*, 85–100. [CrossRef]
15. Lee, S.; Seo, M.H.; Shin, Y. Testing for threshold effects in regression models. *J. Am. Stat. Assoc.* **2011**, *106*, 220–231. [CrossRef]
16. Hansen, B.E. Sample splitting and threshold estimation. *Econometrica* **2000**, *68*, 575–603. [CrossRef]
17. Cho, J.S.; White, H. Testing for regime switching. *Econometrica* **2007**, *75*, 1671–1720. [CrossRef]
18. Newey, W.K.; Powell, J.L. Asymmetric least squares estimation and testing. *Econometrica* **1987**, *55*, 819–847. [CrossRef]
19. Zhang, F.; Li, Q. A continuous threshold expectile model. *Comput. Stat. Data Anal.* **2017**, *116*, 49–66. [CrossRef]
20. Lerman, P.M. Fitting segmented regression models by grid search. *J. R. Stat. Soc. Appl. Stat. Ser. C* **1980**, *29*, 77–84. [CrossRef]

21. Zhou, X.; Zhang, F. A new estimation method for continuous threshold expectile model. *Commun. Stat.-Simul. Comput.* **2018**, *47*, 2486–2498. [CrossRef]
22. Yan, Y.; Zhang, F.; Zhou, X. A note on estimating the bent line quantile regression model. *Comput. Stat.* **2017**, *32*, 611–630. [CrossRef] [PubMed]
23. He, X.; Shao, Q.M. A general bahadur representation of M-Estimators and its application to linear regression. *Ann. Stat.* **1996**, *24*, 2608–2630. [CrossRef]