Computational Statistics
https://doi.org/10.1007/s00180-024-01526-w

ORIGINAL PAPER

Empirical likelihood change point detection in quantile regression models

Suthakaran Ratnasingam[1] · Ramadha D. Piyadi Gamage[2]

Received: 20 April 2023 / Accepted: 1 July 2024
This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2024

Abstract

Quantile regression is an extension of linear regression which estimates a conditional quantile of interest. In this paper, we propose an empirical likelihood-based non-parametric procedure to detect structural changes in the quantile regression models. Further, we have modified the proposed smoothed empirical likelihood-based method using adjusted smoothed empirical likelihood and transformed smoothed empirical likelihood techniques. We have shown that under the null hypothesis, the limiting distribution of the smoothed empirical likelihood ratio test statistic is identical to that of the classical parametric likelihood. Simulations are conducted to investigate the finite sample properties of the proposed methods. Finally, to demonstrate the effectiveness of the proposed method, it is applied to urinary Glycosaminoglycans (GAGs) data to detect structural changes.

1 Introduction

Quantile regression (QR) has gained popularity in recent years due to its appealing advantages over the traditional ordinary least square (OLS) regression model. As an alternative to the ordinary least squares (OLS) regression model, the QR technique was developed by Koenker and Bassett (1978). Unlike OLS models, QR models are used to explain changes in the conditional quantiles of the response variable as a result of changes in the covariates. QR model has been widely used in various areas including biomedical studies, economics, environmental studies, longitudinal studies, and survival analysis, see, for example, Buchinsky (1998), Koenker and Hallock (2001), Koenker and Geling (2001), Koenker (2004), Wei et al. (2006), and many others.

✉ Suthakaran Ratnasingam
suthakaran.ratnasingam@csusb.edu

1 Department of Mathematics, California State University, San Bernardino, San Bernardino, CA 92407, USA

2 Department of Mathematics, Western Washington University, Bellingham, WA 98229, USA

Published online: 10 July 2024

Owen (1988) developed a powerful nonparametric method known as the empirical likelihood (EL) of which the theoretical and practical applications have been extensively studied. EL method has many appealing asymptotic properties. For example, Owen (1990, 2001) showed that the empirical likelihood ratio (ELR) statistic follows an asymptotic chi-square distribution. This is the nonparametric counterpart to the well-known Wilk's theorem. Many researchers investigated the EL-based inference for semiparametric and nonparametric regression models. For instance, the smoothed EL (SEL) method developed by Whang (2006) estimates the parameters and constructs confidence regions for the QR model. Zhao and Chen (2008) proposed EL inference for censored median regression models. A Bayesian empirical likelihood approach to quantile regression was proposed by Yang and He (2012). The empirical likelihood method was used for quantile regression models with applications in censored data by Gao et al. (2021). Tang and Leng (2011) considered the EL for QR in longitudinal data analysis. Wang and Zhu (2011) developed inference procedures in quantile regression using the EL method for longitudinal data. Only a handful of studies have investigated the change point problem based on EL in linear and non-linear models. For example, Liu et al. (2008) studied change point detection procedures in linear regression models using EL inference. Ciuperca and Salloum (2015) proposed an EL-based nonparametric method to detect change points in a general nonlinear model. While Ciuperca and Salloum (2015) primarily focuses on developing empirical likelihood tests in a posteriori change-point nonlinear models, our paper significantly diverges by specifically addressing the extension of change-point detection to quantile regression models using empirical likelihood methods. This shift in focus marks a departure from nonlinear models to quantile regression, thereby addressing a different aspect of statistical modeling and inference. Further, Ciuperca and Salloum (2015) proposed a posteriori change-point nonlinear model. The posteriori change point typically refers to a change point that is detected after observing the data. In other words, it involves analyzing historical data or observed data retrospectively to identify points or times where there is evidence of a change in the underlying statistical properties or characteristics of the data. Also, this can be used in situations where the data is collected continuously or over a period of time. However, the proposed method in this paper does not require real-time data or continuous monitoring. The proposed method can be used when the entire dataset is available upfront ("offline data") or when real-time monitoring is not feasible or necessary.

A constrained maximization in the profile empirical likelihood function necessitates that the convex hull of the estimating equation contains zero vector as an interior point, see, Chen et al. (2008). To solve the issue, an adjusted empirical likelihood (AEL) approach was developed by Chen et al. (2008) which ensures the solution to the maximization problem exists and the same asymptotic optimality features are maintained. Moreover, to address the under-coverage problem in small samples, the transformed empirical likelihood (TEL) was developed by Jing et al. (2017). In this paper, we study the change point problem in quantile regression models using EL-based approaches. To the best of our knowledge, no work has been conducted to investigate the change point problem in quantile regression using the EL approach.

In this paper, we propose an offline SEL-based change point detection procedure to monitor structural change in quantile regression models. The rest of the paper is structured as follows. We provide a brief description of the change point detection method in quantile regression models in Sect. 2. The main theoretical results and two modified SEL procedures are given in Sect. 3. Simulations to examine the finite-sample performance of the suggested procedures and comparison between the proposed methods are conducted in Sect. 4. A real data application is given in Sect. 5. Section 6 provides some discussion. The proofs are deferred to the Appendix.

2 Methodology

Let $\tau \in (0,1)$ be the quantile level of interest. The traditional quantile regression model is given below.

$$y_i = x_i^\top \beta_\tau + e_i, \quad i = 1, \dots, n, \tag{1}$$

where x_i is the ith observation of the covariates X and a p–dimensional design vector, y_i is the ith observation of the response Y, β_τ is a $p \times 1$ vector of unknown parameters, and e_i is the random error satisfying $P(e_i < 0|x_i) = \tau$ for any i. Thus, the quantile regression estimator $\hat{\beta}_\tau$ of β_τ is given by,

$$\min_{\beta \in \mathbb{B}} H_n(\beta) = \arg\min_{\beta \in \mathbb{B}} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau\left(y_i - x_i^\top \beta_\tau\right) \tag{2}$$

where $\mathbb{B}$ is the parameter space and $\rho_\tau(u) = u\{\tau - I(u < 0)\}$ is the quantile loss function. For independent data, Koenker and Bassett (1978) showed that $\hat{\beta}_\tau$ is $n^{1/2}$—consistent and asymptotically normal. Under the above model assumptions, β_τ satisfies the following estimating equation:

$$E\left[x_i \psi(y_i, x_i, \beta_\tau)\right] = 0, \quad i = 1, 2, \dots, n, \tag{3}$$

where $\psi(y, x, \beta_\tau) = I(x^\top \beta_\tau - y > 0) - \tau$ is the quantile score function, and $I(\cdot)$ is the indicator function. Consider $\psi_i(\beta_\tau) = \psi(y_i, x_i, \beta_\tau)$, and $Z(X_i, \beta_\tau) = Z_i(\beta_\tau) = x_i^\top \psi_i(\beta_\tau)$. Let $p_1 \dots, p_n$ be non-negative numbers satisfying $\sum_{i=1}^n p_i = 1$. The quantile empirical log-likelihood ratio function is defined as

$$l(\beta) = -2 \max\left\{ \sum_{i=1}^{n} \log(np_i) \middle| p_i \geq 0, \sum_{i=1}^{n} p_i = 1, \sum_{i=1}^{n} p_i Z_i(\beta_\tau) = 0 \right\}. \tag{4}$$

The Lagrange multiplier method leads to

$$p_i = \frac{1}{n} \frac{1}{1 + \lambda(\beta_\tau)^\top Z_i(\beta_\tau)}, \tag{5}$$

where $\lambda(\beta_\tau)$ is p–dimensional Lagrange multiplier satisfying

$$\frac{1}{n}\sum_{i=1}^{n}\frac{Z_i(\beta_\tau)}{1+\lambda(\beta_\tau)^\top Z_i(\beta_\tau)}=0. \tag{6}$$

Thus, the empirical log-likelihood ratio statistic can be written as

$$l(\beta_\tau)=2\sum_{i=1}^{n}\log\left[1+\lambda(\beta_\tau)^\top Z_i(\beta_\tau)\right] \tag{7}$$

with $\lambda(\beta_\tau)$ satisfying (6). Further, the Eq. (6) can be solved by the modified Newton–Raphson algorithm of Chen and Gupta (2000). Thus, the maximum empirical likelihood estimator of β_τ is defined as,

$$\hat{\beta}_{\tau EL}=\arg\min_{\beta_\tau\in\mathbb{B}} l\big(\beta_\tau\big) \tag{8}$$

To improve the accuracy of empirical likelihood, we can use higher-order refinements through Taylor series approximation, but this approach relies on smooth moment restrictions. As pointed out by Whang (2006), unfortunately, the quantile score function $\psi(\cdot)$ is not differentiable at β when $y=x^\top\beta_\tau$. To address this issue and achieve higher-order accuracy, we can use a modified version of empirical likelihood called smooth empirical likelihood (SEL) proposed by Whang (2006). To approximate the quantile score function $\psi(\cdot)$ with a smooth function $\psi_h(\cdot)$, we use a kernel function $K(\cdot)$ and define $G(x)=\int_{-\infty}^{x}K(u)du$. We also introduce a positive bandwidth parameter h and define $G_h(x)=G(x/h)$. Then, we approximate $\psi(\cdot)$ by $\psi_h(y_i,x_i,\beta_\tau)=G_h(x_i^\top\beta_\tau-y_i)-\tau$. Finally, we define $Z_{hi}(\beta_\tau)=x_i^\top\psi_{hi}(\beta_\tau)$. The quantile smooth empirical log-likelihood ratio function is defined as

$$l_h(\beta_\tau)=-2\max\left\{\sum_{i=1}^{n}\log(np_i)\middle|p_i\geq 0,\sum_{i=1}^{n}p_i=1,\sum_{i=1}^{n}p_iZ_{hi}(\beta_\tau)=0\right\}. \tag{9}$$

Thus, the maximum smooth empirical likelihood estimator of β_τ is defined as,

$$\hat{\beta}_{\tau SEL}=\arg\min_{\beta_\tau\in\mathbb{B}} l_h(\beta_\tau) \tag{10}$$

Whang (2006) showed that under some regularity conditions, $\hat{\beta}_{\tau EL}$ and $\hat{\beta}_{\tau SEL}$ have the same asymptotic distribution as h goes to zero sufficiently fast when $n\longrightarrow\infty$.

2.1 Change point problem

Consider a change point model with at most one change at an unknown change point location k. Thus, the change point problem for the quantile regression model can be expressed as follows:

$$y_i=\begin{cases} x_i^\top\beta_\tau^1+e_i & i=1,\dots,k\\ x_i^\top\beta_\tau^2+e_i & i=(k+1),\dots,n\end{cases} \tag{11}$$

where $k \in \{2, \dots, (n-1)\}$ is the unknown change point location. Let $\beta_\tau^1 = (\beta_{11}, \dots, \beta_{1p})^\top$ and $\beta_\tau^2 = (\beta_{21}, \dots, \beta_{2p})^\top$. Let $r \geq 2$ be an integer. We denote $F(u_i|x)$ as the marginal distribution function of e_i conditional on $X_i = x$ and and $f(u_i|x)$ as the marginal density of e_i with respect to the Lebesgue measure. Further, $S = E\left[x^\top f(0|x)x\right]$, and $\Sigma = E\left[x_i^\top \psi_i(\beta_\tau)\psi_i(\beta_\tau)^\top x_i\right]$, where $\psi_i(\beta_\tau) = \psi(y_i, x_i, \beta_\tau)$. The testing hypothesis can be defined as,

$$H_0 : \beta_\tau^1 = \beta_\tau^2 \quad vs \quad H_1 : \quad \beta_\tau^1 \neq \beta_\tau^2$$

In particular, our null hypothesis implies that there is no change in the regression coefficient versus the alternative that there is at most one change in the regression coefficient. According to Owen (1990), the Eq. (9) holds when there is no change point in the regression coefficient.

Let $p = (p_1, \dots, p_k)^\top$ and $q = (p_{k+1}, \dots, p_n)^\top$ be probability vectors such that $\sum_{i=1}^k p_i = 1, \sum_{j=k+1}^n q_j = 1$ and $p_i, q_j \geq 0$. Under H_0, the smooth empirical log-likelihood (SEL) for β_τ is,

$$\Lambda_{H_0,k} = \sup_{H_0} \left\{ \sum_{i=1}^k \log(kp_i) + \sum_{j=k+1}^n \log((n-k)q_j) \;\middle|\; \sum_{i=1}^k p_i Z_{hi}(\beta_\tau) = \sum_{j=k+1}^n q_j Z_{hj}(\beta_\tau) = 0 \right\}. \tag{12}$$

Similarly, under the alternative hypothesis, the SEL is

$$\Lambda_{H_1,k} = \sup_{H_1} \left\{ \sum_{i=1}^k \log(kp_i) + \sum_{j=k+1}^n \log((n-k)q_j) \;\middle|\; \sum_{i=1}^k p_i Z_{hi}(\beta_\tau^1) = 0, \sum_{j=k+1}^n q_j Z_{hj}(\beta_\tau^2) = 0 \right\}. \tag{13}$$

Suppose the change occurs in the regression coefficient at a fixed location k. Let Λ_k be the ratio between $\Lambda_{H_0,k}$ and $\Lambda_{H_1,k}$. Then the smooth empirical likelihood ratio (SELR) statistic can be defined as,

$$\begin{aligned} -2\log\Lambda_k &= -2\log \frac{\sup_{H_0}\left\{ \prod_{i=1}^k p_i \prod_{j=k+1}^n q_j \middle| \sum_{i=1}^k p_i Z_{hi}(\beta_\tau) = \sum_{j=k+1}^n q_j Z_{hj}(\beta_\tau) = 0 \right\}}{\sup_{H_1}\left\{ \prod_{i=1}^k p_i \prod_{j=k+1}^n q_j \middle| \sum_{i=1}^k p_i Z_{hi}(\beta_\tau^1) = 0, \sum_{j=k+1}^n q_j Z_{hj}(\beta_\tau^2) = 0 \right\}} \\ &= \Lambda_{H_0,k} - \Lambda_{H_1,k}. \end{aligned} \tag{14}$$

The null hypothesis is rejected for large value of $\max_{1 \leq k \leq n} -2\log\Lambda_k$. Under H_1, the SELR has an asymptotic χ^2 distribution. Take $\theta_{nk} = \frac{k}{n}$. The Lagrange multiplier method provides,

$$p_i = \frac{1}{n\theta_{nk}\left[1 + \theta_{nk}^{-1}\lambda_1^\top Z_{hi}(\beta_\tau^1)\right]} \quad \text{and} \quad q_i = \frac{1}{n\left(1-\theta_{nk}\right)\left[1-(1-\theta_{nk})^{-1}\lambda_2^\top Z_{hj}(\beta_\tau^2)\right]},$$

where λ_1 and λ_2 are chosen such that $\sum_{i=1}^{k} p_i Z_{hi}\left(\beta_\tau^1\right) = 0,$ and $\sum_{j=k+1}^{n} q_j Z_{hj}\left(\beta_\tau^2\right) = 0,$ respectively. Thus, under H_0 we have,

$$\Lambda_{H_0,k} = 2 \inf_{\beta} \sup_{\lambda_1,\lambda_2} \left\{ \sum_{i=1}^{k} \log(1 + \theta_{nk}^{-1}\lambda_1^\top Z_{hi}(\beta_\tau^1)) + \sum_{j=k+1}^{n} \log(1-(1-\theta_{nk})^{-1}\lambda_2^\top Z_{hj}(\beta_\tau^1)) \right\} \tag{15}$$

Then the empirical likelihood ratio statistic is defined in (14) can be rewritten as,

$$-2\log\Lambda_k = 2\left\{l_{SEL}(\tilde{\beta}, 0) - l_{SEL}(\tilde{\beta}, \tilde{\delta})\right\} \tag{16}$$

where $\tilde{\beta}$ is the estimator of β_τ and $\tilde{\delta} = \tilde{\beta}_\tau^1 - \tilde{\beta}_\tau^2$ is the estimator of $\delta = \beta_\tau^1 - \beta_\tau^2$. Further, $l_{SEL}(\tilde{\beta}, 0)$ and $l_{SEL}(\tilde{\beta}, \tilde{\delta})$ are the smoothed empirical log-likelihood obtained under the null and alternative hypotheses, respectively. Since k is unknown, H_0 is rejected when the maximally selected log-likelihood ratio statistic,

$$Z_n = \max_{\theta_{nk}\in\Theta_n^*} \left\{-2\log\Lambda_k\right\},$$

where $\Theta_n^* = \{k/n; k = 1, \dots, n\}$, is sufficiently large. For small values of k or $n-k$, then the minimax estimators of empirical likelihood $(\tilde{\beta}, \tilde{\lambda})$ may not exist. As a result, we consider the range of k to be chosen at random which in return gives a trimmed version of the likelihood ratio statistic. The range of k is chosen at random as follows. The trimmed log-likelihood ratio statistic is defined as,

$$\tilde{Z}_n = \max_{\theta_{nk}\in\Theta_n} \left\{-2\log\Lambda_k\right\}, \tag{17}$$

where $\Theta_n = \{k/n; k = n_{T_1}, n_{T_1}+1, \dots, n-n_{T_2}\}$. The selection of n_{T_1} and n_{T_2} is arbitrary. In our work, we choose $n_{T_1} = n_{T_1} = 2[n^{1/2}]$, where $[x]$ represents the largest integer not larger than x. Under H_0, the test statistic $\tilde{Z}_n$ follows an asymptotic extreme value distribution. Since asymptotic tests frequently tend to be overly conservative in finite samples, the convergence to the extreme value limit can be slow. For the asymptotic null distribution to follow the extreme distribution, the selection of n_{T_1} and n_{T_2} must meet the requirement stated by Csörgö and Horváth (1997) (Ref. p. 366, Theorem A.3.4). Any combination of n_{T_1} and n_{T_2} would be effective so long as this requirement is met.

We propose two modified empirical likelihood-based approaches, AEL and TEL, to improve the robustness and accuracy of structural change detection in these models. AEL tackles the challenge posed by the convex hull problem in traditional empirical likelihood methods. By strategically introducing an artificial data point into the dataset, AEL ensures that zero lies within the convex hull, improving the reliability of likelihood estimation. This method is particularly valuable in scenarios with small sample sizes or complex data distributions, where traditional approaches may be inefficient. The TEL, on the other hand, offers a simpler yet effective strategy to enhance coverage probabilities in empirical likelihood methods. The TEL uses the truncated quadratic transformation of the original empirical log-likelihood ratio. Through this straightforward transformation, TEL maintains the same properties of empirical likelihood while improving its performance.

3 Main results

In this section, we establish the asymptotic properties of the test statistic. Let $r \geq 2$ be an integer. To establish the asymptotic null distribution of the test statistics $\tilde{Z}_n$, we need the following assumptions. Note that these assumptions are the ones typically used for achieving smooth QR estimation.

A1. $\{(y_i, x_i) : i = 1, \dots, n\}$ are i.i.d, random vectors
A2. The support of x, $\mathcal{X}$, is compact,
A3. The sequence $\{x_i, 1 \leq i < \infty\}$ is strictly stationary with a finite fourth moment, that is, $\lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=1}^{n} ||x_i||^4 < \infty$ a.s.
A4. The parameter vector β_τ is an interior point of the parameter space $\mathbb{B}$, a compact subset of R^p.
A5. (a) $F(0|x) = q$ for almost every $x \in \mathcal{X}$
  (b) For all u in a neighborhood of 0 and almost every x, $f(u|x)$ exists, is bounded away from zero, and is r times continuously differentiable with respect to u.

A6. S and Σ are nonsingular.
A7. The positive bandwidth parameter h satisfies (a) $nh^{2r} \longrightarrow 0$ and (b) $nh/\log n \longrightarrow \infty$ as $n \longrightarrow \infty$.
A8. (a) $K(\cdot)$ is bounded and compactly supported on $[-1, 1]$.
  (b) For some constant $C_K \neq 0$, $K(\cdot)$ is an rth - order kernel, i.e., $\int u^j K(u)du = 1$ if $j = 0$; of if $1 \leq j \leq r-1$; C_K if $j = r$
  (c) Let $\widetilde{G}(u) = \left(G(u), G^2(u), \dots, G^{L+1}(u)\right)^\top$ for some $L \geq 1$, where $G(u) = \int_{v<u} K(v)dv$. For any $\theta \in R^{L+1}$ satisfying $||\theta|| = 1$, there is a partition of $[-1, 1]$, $-1 = a_0 < a_1 < \cdots < a_{L+1} = 1$ such that $\theta^\top \widetilde{G}(u)$ is either strictly positive or strictly negative on (a_{l-1}, a_l) for $l = 1, \dots, L+1$.

Assumptions A1–A3 are standard for linear models, see, for example, Wang and Zhu (2011) and Liu et al. (2008). Assumptions A4–A8 are specific to the quantile regression models which are considered in Whang (2006) and Wang and Zhu (2011). Further, assumptions A1, A2, and A5–A8 are typically used to establish the asymptotic properties of the smoothed empirical likelihood procedure for quantile regression estimations. Let $1 < m < n$ and $\lambda = \left(\lambda_1^\top, \lambda_2^\top\right)^\top$. Also, let $\psi_{hi}(\beta_\tau) = \psi_h(y_i, x_i, \beta_\tau)$, and $Z_h(X_i, \beta_\tau) = Z_{hi}(\beta_\tau) = x_i^\top \psi_{hi}(\beta_\tau)$. We define the score functions, $S_1(\beta_\tau, \lambda)$ and $S_2(\beta_\tau, \lambda)$, as below.

$$S_1(\beta_\tau, \lambda) = \frac{\partial \Lambda_{H_0,k}}{\partial \lambda} = \frac{1}{n} \sum_{m_1} \frac{1}{(1 + \theta_{m_1}^{-1} \lambda^\top Z_{hm_1}(\beta_\tau))} \theta_{m_1}^{-1} Z_{hm_1}(\beta_\tau), \quad \text{and}$$

$$S_2(\beta_\tau, \lambda) = \frac{\partial \Lambda_{H_0,k}}{\partial \beta_\tau} = \frac{1}{n} \sum_{m_2} \frac{1}{(1 + \theta_{m_2}^{-1} \lambda^\top Z_{hm_2}(\beta_\tau))} \theta_{m_2}^{-1} \left(\frac{\partial Z_{hm_2}(\beta_\tau)}{\partial \beta_\tau} \right)^\top Z_{hm_2}(\beta_\tau)$$

where

$$\frac{1}{\theta_m} = \frac{1}{\theta_{nk}} I(1 < m \leq k) + \frac{1}{(1 - \theta_{nk})} I(k + 1 \leq m < n)$$

and

$$Z_{hm}(\beta_\tau) = x_i^\top \psi_h(y_i, x_i, \beta_\tau) I(1 < m \leq k) + x_j^\top \psi_h(y_j, x_j, \beta_\tau) I(k + 1 \leq m < n)$$

where $i = 1, \ldots, k$ and $j = k + 1, \ldots, n$. It should be noted that for $i = 1, \ldots, k$, $E\left[x_i^\top \psi_h(y_i, x_i, \beta_\tau)\right] = 0$ and for $j = k + 1, \ldots, n$, $E\left[x_j^\top \psi_h(y_j, x_j, \beta_\tau)\right] = 0$. Under certain regularity conditions, Qin and Lawless (1994) showed, there exists $(\tilde{\beta}, \tilde{\lambda})$ such that,

$$S_1(\tilde{\beta}, \tilde{\lambda}) = 0 \ \text{ and } \ S_2(\tilde{\beta}, \tilde{\lambda}) = 0.$$

The following theorem pertains to the asymptotic null distribution of the test statistic $\tilde{Z}_n$.

Theorem 3.1 *Suppose that assumptions A1-A8 are satisfied and* $E||Z_{hi}(\beta_\tau))||^3 < \infty, E||x||^4 < \infty$ *and* $E(Z_{hi}(\beta_\tau) Z_{hi}(\beta_\tau)^\top)$ *is nonsingular. If* H_0 *is true, then we have*

$$\lim_{n \to \infty} Pr\left\{ A(\log u(n)) (\tilde{Z}_n)^{\frac{1}{2}} \leq t + D_s(\log u(n)) \right\} = \exp(-e^{-t})$$

for all t, where $A(x) = (2 \log x)^{\frac{1}{2}}, D_s(x) = 2 \log x + \left(\frac{s}{2}\right) \log \log x - \log \Gamma\big(s/2\big)$,

$$u(n) = \frac{n^2 + \left(2\left\lfloor n^{\frac{1}{2}} \right\rfloor\right)^2 - 2n\left\lfloor n^{\frac{1}{2}} \right\rfloor}{\left(2\left\lfloor n^{\frac{1}{2}} \right\rfloor\right)^2},$$

and s is the dimension of the parameter δ.

It should be noted that the definition of $D_s(\cdot)$ is similar to the equation (1.3.2) in Csörgö and Horváth (1997). In our case, the dimension $s = 2$ cancels out the 3rd term in the equation.

Theorem 3.2 *Suppose that assumptions A1-A8 are satisfied, under the alternative hypothesis, if* $\theta_0 \in (0,1)$ *and* $\lim_{n\longrightarrow\infty}\theta_{nk} = \theta$ *, then* $-2\log\Lambda_k \longrightarrow \omega(\theta)$ *almost surely as* $n \longrightarrow \infty$*, where* $\omega(\theta)$ *is strictly increasing on* $(0,\theta_0)$*, decreasing on* $(\theta_0,1)$ *and* $\max_{0\le\theta\le 1}\omega(\theta) = \omega(\theta_0)$.

Theorem 3.3 *Suppose that assumptions A1-A8 are satisfied. Under the alternative hypothesis, if* $\theta_0 \in (0,1)$ *and* $\lim_{n\longrightarrow\infty}\theta_{nk} = \theta$ *, there exists a constant* $c > 0$ *such that* $P(\tilde{Z}_n > cn) \longrightarrow 1$*, that is, the test statistic* $\tilde{Z}_n$ *is consistent.*

Sketch of these proofs are given in the Appendix.

As discussed earlier, the EL method needs the convex hull of the estimating equation to contain a zero as an interior point. In the event that there is no solution, Owen (2001) proposed assigning $-\infty$ to the empirical log-likelihood ratio statistic. However, this approach frequently has drawbacks. Chen et al. (2008) proposed adding a pseudo term to ensure that the zero-vector is within the convex hull. Let $Z_{hi} = Z_{hi}(\beta_\tau)$ for $i = 1,\dots,n$. Reprising Chen et al. (2008), for any given β_τ and some positive constant a_n. Let $\bar{Z}_n^* = \bar{Z}_n^*(\beta_\tau) = \frac{1}{n}\sum_{i=1}^n Z_{hi}$ for any given β_τ. We define an additional term, Z_{n+1}^* as,

$$Z_{n+1}^* = Z_{n+1}^*(\beta_\tau) = -\frac{a_n}{n}\sum_{i=1}^{n} Z_{hi}(\beta_\tau) = -a_n\bar{Z}_n^*. \tag{18}$$

The original empirical log-likelihood ratio $l_h(\beta_\tau)$ defined in (9) can be adjusted as

$$l_h^*(\beta_\tau) = \max\left\{\sum_{i=1}^{n+1}\log((n+1)p_i)\middle| p_i \ge 0, \sum_{i=1}^{n+1} p_i = 1, \sum_{i=1}^{n+1} p_i Z_{hi}(\beta_\tau) = 0\right\}. \tag{19}$$

Chen et al. (2008) showed that, as $n \longrightarrow \infty$, $-2l_h^*(\beta_\tau) \longrightarrow \chi^2(p)$ in distribution with $a_n = o_p(n^{2/3})$, where p is the dimension of the vector (x_{ij}).

Another issue in the EL method is the under-coverage problem, which is caused by the convex hull constraint, where the coverage probability is typically lower than the nominal level, see Tsao (2013). As a solution, Jing et al. (2017) developed a transformed version of the EL which is theoretically and computationally simple, and preserves the properties P1–P4 given in Jing et al. (2017). Piyadi Gamage and Ning (2020) extended EL-based approach to transformed EL ratio test statistic for autoregressive fractionally integrated moving average (ARFIMA) model. Thus, we propose the transformed SEL ratio for the quantile regression model, $l_h^T(\beta_\tau)$ is

$$l_h^T(l_h(\beta_\tau),\gamma) = l_h(\beta_\tau)\times\max\left\{1-\frac{l_h(\beta_\tau)}{n}, 1-\gamma\right\}, \tag{20}$$

where $\gamma \in [0,1]$. Jing et al. (2017) sets $\gamma = 0.5$ by default and $l_h(\beta_\tau)$ defined in eq (9). Similar to Piyadi Gamage and Ning (2020), it can be shown that the TSEL in (20) satisfies the four conditions given below. See Piyadi Gamage and Ning (2020) for more details.

1. For any β, $0 \le l_h^T(\beta_\tau) \le l_h(\beta_\tau)$,
2. $l_h^T(\beta_1) \le l_h^T(\beta_2)$ if and only if $l_h(\beta_1) \le l_h(\beta_2)$ for any $\beta_1, \beta_2 \in \mathbb{B}$,
3. Limiting distribution of $l_h^T(\beta_\tau)$ is same as that of $l_h(\beta_\tau)$, that is $l_h^T(\beta_\tau) = l_h(\beta_\tau) + o_p(1)$ and
4. $l_h^T(\beta_\tau)$ contours have data driven shape and they are centered around $\tilde{\beta}$, where $\tilde{\beta}$ is the estimator of β_τ.

It can be shown that the asymptotic distribution of $-2l_h^T(\beta_\tau)$ is χ_p^2 since $l_h^T(\beta_\tau)$ satisfies the four conditions above and $-2l_h(\beta_\tau) \longrightarrow \chi^2(p)$, as $n \to \infty$.

4 Simulation study

In this section, we conduct a simulation study to assess the performance of the proposed SEL-based change point detection method. Our change point model is given in (11). Without loss of generality, we consider that $p = 1$. Hence, in our simulation study, the pre-change parameter $\beta_\tau^1 = 2$ and the post-change parameter $\beta_\tau^2 = 3$. Under the null hypothesis, the model errors, e_i are generated from $N(0,\ 1)$. Under the alternative hypothesis, the model errors, e_i are generated from the following distributions.

- Case 1 (Homoskedastic errors): $e_i \sim N(0, 1)$.
- Case 2 (Heavy-tailed errors): $e_i \sim \text{Cauchy}(0, 2)$.
- Case 3 (Skewed errors): $e_i \sim SN(0, 1, 3)$.

We also consider various sample sizes, n, ranging from small to large, as well as different change point locations. First, we computed the empirical critical values under different significance levels via simulations. The steps are given below.

Step 1: Under null hypothesis, we generate data from $N(0,\ 1)$ with various sample sizes $n = \{50, 100, 150, 300\}$ and quantile level $\tau = 0.25, 0.5, 0.75$.

Step 2: For each generated sample, for example, the SEL-based method we calculate the test statistic $\tilde{Z}_n$ given in equation (17). The test statistics for the ASEL and TSEL methods can also be calculated similarly.

Step 3: We repeat the above steps $B = 1000$ times. Then the critical value is the $100(1-\alpha)$th quantile of the asymptotic distribution obtained in Step 2, where the significance level $\alpha = 0.05$.

The SEL-based methods proposed in this study depend on the smoothing parameter h. As indicated by Wang and Zhu (2011), the SEL demonstrates robustness to bandwidth variations and provides satisfactory coverage probabilities for $h \in \left[n^{-0.5}, n^{-0.9}\right]$. Therefore, in our simulation study, we adopt $n^{-0.8}$. We computed the empirical critical values using the bootstrap method because the underlying distribution is asymptotic. It is noteworthy that the empirical critical values were found to be more similar to the theoretical critical values for the corresponding significance levels. Therefore, we used the empirical critical values for power calculations. This approach was selected based on the better agreement between the empirical and theoretical values. Moreover, in the real data application, the sample size of the data is large enough to guarantee the convergence of the test statistic to the extreme value distribution. Hence, the theoretical critical values derived from Theorem 1 given in Table 1) are used to perform the real data application.

First, we assess the empirical type I error for all three proposed methods. The results are summarized in Table 2. We considered different sample sizes $n = 50, 100, 150, 300$ to represent small, medium & large samples at different quantile levels, $\tau = 0.25, 0.5, 0.75$ under various settings. For instance, consider the Type I error rates for $\tau = 0.25$ and method SEL. As the sample size increases from 50 to 300, the Type I error rates decrease: from 0.052 to 0.030. This trend is also noticeable across different true effect sizes and methods. Larger sample sizes generally contribute to more accurate detection and improved control of Type I errors. When the quantile level (τ) is set at 0.25 and the sample size (n) is 50, the TSEL method demonstrates superior performance in controlling Type I errors. Specifically, it yields Type I error rates of 0.045, 0.062, and 0.066 for cases (distributions) 1 to 3, respectively. This illustrates the effectiveness of the TSEL approach in comparison to the SEL and ASEL methods in controlling Type I errors under different error distributions. Moreover, when $\tau = 0.5$ or 0.75, The ASEL and TSEL methods exhibit type I error rates that closely approximate the nominal level, $\alpha = 0.05$, and reflect comparable performance. However, the SEL method tends to yield a slightly higher type I error than the nominal level. The TSEL method shows robustness in effectively controlling type I errors across diverse error distributions, including scenarios with Homoskedastic errors (case 1), Heavy-tailed errors (case 2), and Skewed errors (case 3). This is particularly important in detecting shifts in quantile regression and shows its reliability as a tool for identifying meaningful changes in the distribution of the data, even in the presence of different error structures. Hence, if the sample size is $n < 100$ and the quantile level $\tau < 0.5$, we suggest using the TSEL method. All three proposed methods are suitable for large sample sizes such as $n \geq 100$. The type I errors are graphed in Fig. 1.

We have conducted power analysis with different sample sizes including $n = 50, 100, 150, 300$ at multiple change point locations with different quantile levels $\tau = 0.25, 0.5, 0.75$. The results for the proposed three methods are summarized in Tables 3, 4, and 5. It can be seen that the power tends to increase as the sample size increases. Furthermore, when the change location is closer to the middle of the data set, the empirical power tends to be higher as compared to change locations closer to the beginning of the data set. This pattern is consistent with the different error distributions we have considered. Moreover, the empirical power is higher for the $N(0, 1)$

Table 1 Theoretical critical values

$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 1\%$
2.2504	2.9702	4.6001

Table 2 Empirical type I error with nominal level $\alpha = 0.05$

τ	n	Method	$N(0, 1)$	Cauchy(0, 2)	$SN(0, 1, 3)$
0.25	50	SEL	0.052	0.081	0.083
		ASEL	0.047	0.072	0.078
		TSEL	0.045	0.062	0.066
	100	SEL	0.044	0.056	0.057
		ASEL	0.041	0.054	0.055
		TSEL	0.041	0.048	0.054
	150	SEL	0.038	0.044	0.046
		ASEL	0.036	0.038	0.038
		TSEL	0.031	0.036	0.037
	300	SEL	0.030	0.033	0.035
		ASEL	0.029	0.032	0.034
		TSEL	0.028	0.031	0.032
0.5	50	SEL	0.050	0.072	0.073
		ASEL	0.047	0.067	0.069
		TSEL	0.045	0.059	0.062
	100	SEL	0.043	0.055	0.056
		ASEL	0.040	0.049	0.054
		TSEL	0.037	0.045	0.053
	150	SEL	0.035	0.040	0.038
		ASEL	0.032	0.037	0.037
		TSEL	0.030	0.036	0.037
	300	SEL	0.029	0.033	0.035
		ASEL	0.029	0.032	0.034
		TSEL	0.028	0.030	0.032
0.75	50	SEL	0.049	0.061	0.063
		ASEL	0.045	0.065	0.068
		TSEL	0.045	0.060	0.058
	100	SEL	0.041	0.055	0.056
		ASEL	0.039	0.048	0.054
		TSEL	0.036	0.045	0.051
	150	SEL	0.034	0.038	0.038
		ASEL	0.031	0.036	0.037
		TSEL	0.030	0.034	0.036
	300	SEL	0.029	0.033	0.034
		ASEL	0.028	0.031	0.033
		TSEL	0.028	0.030	0.031

error in comparison to the other two error distributions. However, when the sample size increases, the empirical power approaches the nominal value, as expected. Thus, our proposed method is robust for a wide range of scenarios.

Additionally, we conducted the simulations to ASEL and TSEL-based methods. Similar to the SEL-based simulations, we have computed power values for different

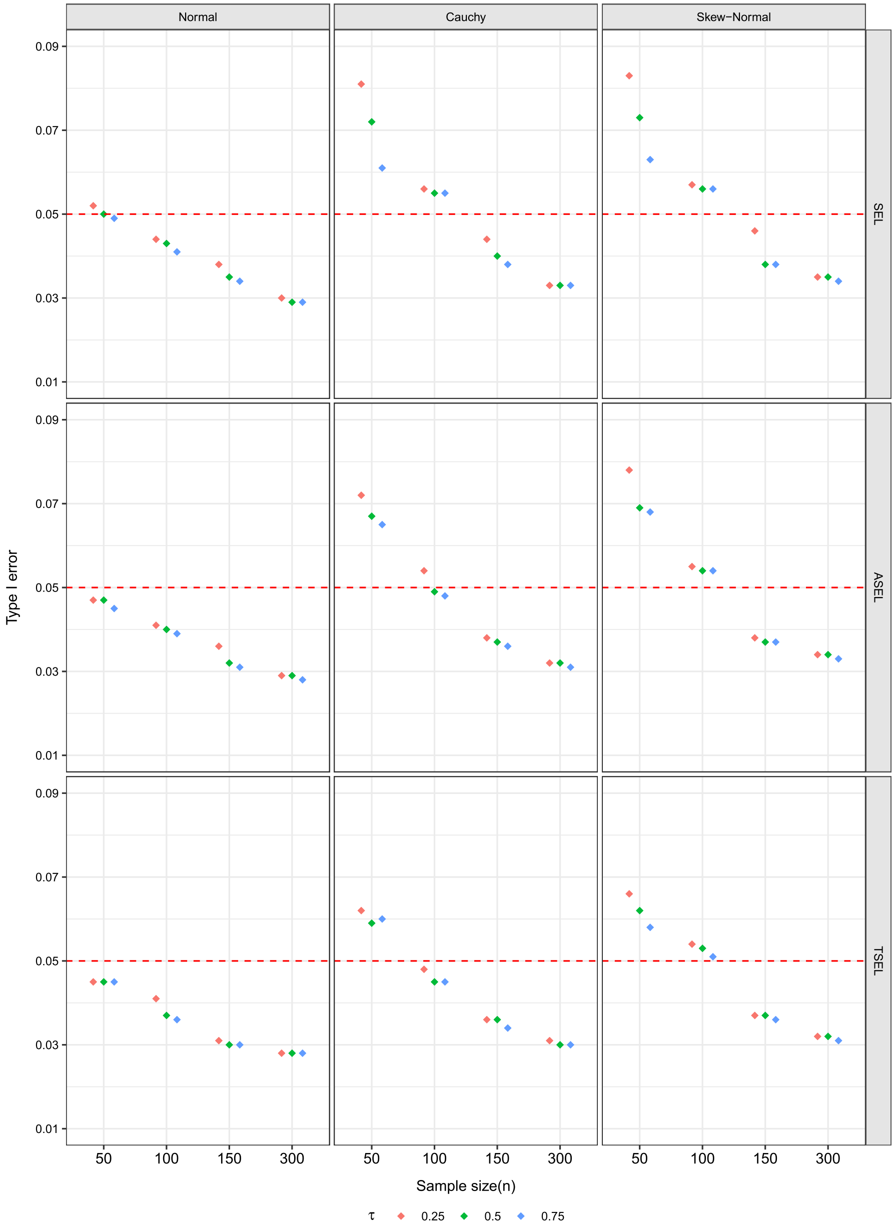

Fig. 1 Empirical comparisons of Type I errors across various distributions, quantile levels, and different sample sizes ($n = 50, 100, 150, 300$) at a nominal level of $\alpha = 0.05$

sample sizes at various true change locations using the ASEL and TSEL methods. Since the test statistics given in (19) and (20) are slightly different, the critical values need to be reassessed. Hence, we used the empirical critical values corresponding to

Table 3 SEL based power values for different sample sizes and error distributions

τ	n	k	$N(0, 1)$	Cauchy(0, 2)	$SN(0, 1, 3)$
0.25	50	15	0.234	0.223	0.199
		20	0.240	0.223	0.201
		25	0.242	0.229	0.203
	100	20	0.343	0.343	0.307
		35	0.349	0.345	0.307
		50	0.353	0.351	0.317
	150	25	0.530	0.533	0.525
		50	0.536	0.549	0.535
		75	0.540	0.547	0.539
	300	75	0.830	0.835	0.799
		100	0.830	0.839	0.803
		150	0.838	0.841	0.815
0.5	50	15	0.342	0.340	0.308
		20	0.374	0.350	0.324
		25	0.388	0.362	0.362
	100	20	0.476	0.466	0.480
		35	0.516	0.496	0.482
		50	0.556	0.532	0.532
	150	25	0.636	0.614	0.608
		50	0.664	0.622	0.638
		75	0.692	0.648	0.640
	300	75	0.904	0.902	0.906
		100	0.926	0.914	0.924
		150	0.940	0.926	0.930
0.75	50	15	0.601	0.596	0.573
		20	0.611	0.598	0.579
		25	0.627	0.600	0.581
	100	20	0.712	0.701	0.704
		35	0.714	0.707	0.706
		50	0.720	0.711	0.714
	150	25	0.856	0.852	0.849
		50	0.862	0.855	0.852
		75	0.866	0.857	0.856
	300	75	0.949	0.948	0.945
		100	0.953	0.949	0.947
		150	0.957	0.955	0.952

each method for power calculations. Simulation results based on ASEL and TSEL are summarized in Tables 4 and 5 respectively. While ASEL-based test statistic performs better than the SEL-based method, the TSEL-based test statistics outperforms both these methods. The reason for this improvement in the power of the test is that the TSEL-based test statistic guarantees that the convex hull contains a zero as an

Table 4 ASEL based power values for different sample sizes and error distributions

τ	n	k	$N(0, 1)$	Cauchy(0, 2)	$SN(0, 1, 3)$
0.25	50	15	0.255	0.243	0.228
		20	0.261	0.247	0.230
		25	0.263	0.249	0.232
	100	20	0.409	0.388	0.377
		35	0.414	0.395	0.376
		50	0.418	0.416	0.387
	150	25	0.629	0.608	0.599
		50	0.635	0.614	0.600
		75	0.639	0.612	0.604
	300	75	0.876	0.869	0.857
		100	0.887	0.873	0.867
		150	0.895	0.875	0.869
0.5	50	15	0.362	0.352	0.312
		20	0.376	0.360	0.328
		25	0.406	0.362	0.366
	100	20	0.612	0.580	0.582
		35	0.632	0.594	0.584
		50	0.636	0.622	0.604
	150	25	0.746	0.734	0.788
		50	0.780	0.756	0.802
		75	0.790	0.782	0.806
	300	75	0.902	0.912	0.940
		100	0.912	0.892	0.958
		150	0.921	0.902	0.960
0.75	50	15	0.625	0.596	0.573
		20	0.636	0.598	0.579
		25	0.652	0.600	0.581
	100	20	0.724	0.721	0.724
		35	0.728	0.725	0.726
		50	0.730	0.729	0.728
	150	25	0.892	0.852	0.849
		50	0.895	0.855	0.852
		75	0.905	0.857	0.856
	300	75	0.973	0.970	0.968
		100	0.977	0.974	0.970
		150	0.981	0.978	0.975

interior point as well as addresses the under-coverage problem, as explained in Section 3.2. For small sample sizes, for example, $n = 50$, the TSEL-based method produces a good power closer to 1 at various change point locations. The SEL method performs worst among all three methods. However, for large n, all three methods perform equally well. In particular, the simulation results indicate that our proposed

Table 5 TSEL based power values for different sample sizes and error distributions

τ	n	k	$N(0, 1)$	Cauchy(0, 2)	$SN(0, 1, 3)$
0.25	50	15	0.407	0.388	0.382
		20	0.415	0.397	0.388
		25	0.418	0.408	0.392
	100	20	0.562	0.548	0.537
		35	0.568	0.555	0.546
		50	0.572	0.561	0.557
	150	25	0.784	0.763	0.759
		50	0.792	0.780	0.768
		75	0.799	0.782	0.782
	300	75	0.896	0.894	0.893
		100	0.914	0.902	0.896
		150	0.925	0.915	0.911
0.5	50	15	0.550	0.528	0.526
		20	0.568	0.546	0.538
		25	0.592	0.574	0.572
	100	20	0.692	0.652	0.642
		35	0.696	0.698	0.668
		50	0.722	0.726	0.688
	150	25	0.836	0.834	0.826
		50	0.860	0.866	0.858
		75	0.868	0.864	0.860
	300	75	0.924	0.920	0.912
		100	0.928	0.924	0.921
		150	0.932	0.926	0.924
0.75	50	15	0.660	0.631	0.608
		20	0.675	0.637	0.618
		25	0.710	0.658	0.639
	100	20	0.794	0.761	0.764
		35	0.817	0.786	0.785
		50	0.830	0.799	0.802
	150	25	0.918	0.898	0.895
		50	0.923	0.883	0.880
		75	0.929	0.881	0.880
	300	75	0.979	0.974	0.974
		100	0.983	0.975	0.976
		150	0.990	0.984	0.984

method is robust under various settings. According to the simulation, the TSEL method is recommended for small sizes. However, all three methods are suitable for large sample sizes. Figure 2 shows the power comparison based on SEL, ASEL, and TSEL methods. Furthermore, it should be noted that power tends to be higher as the change point location gets closer to the center of the dataset. Additionally, when the

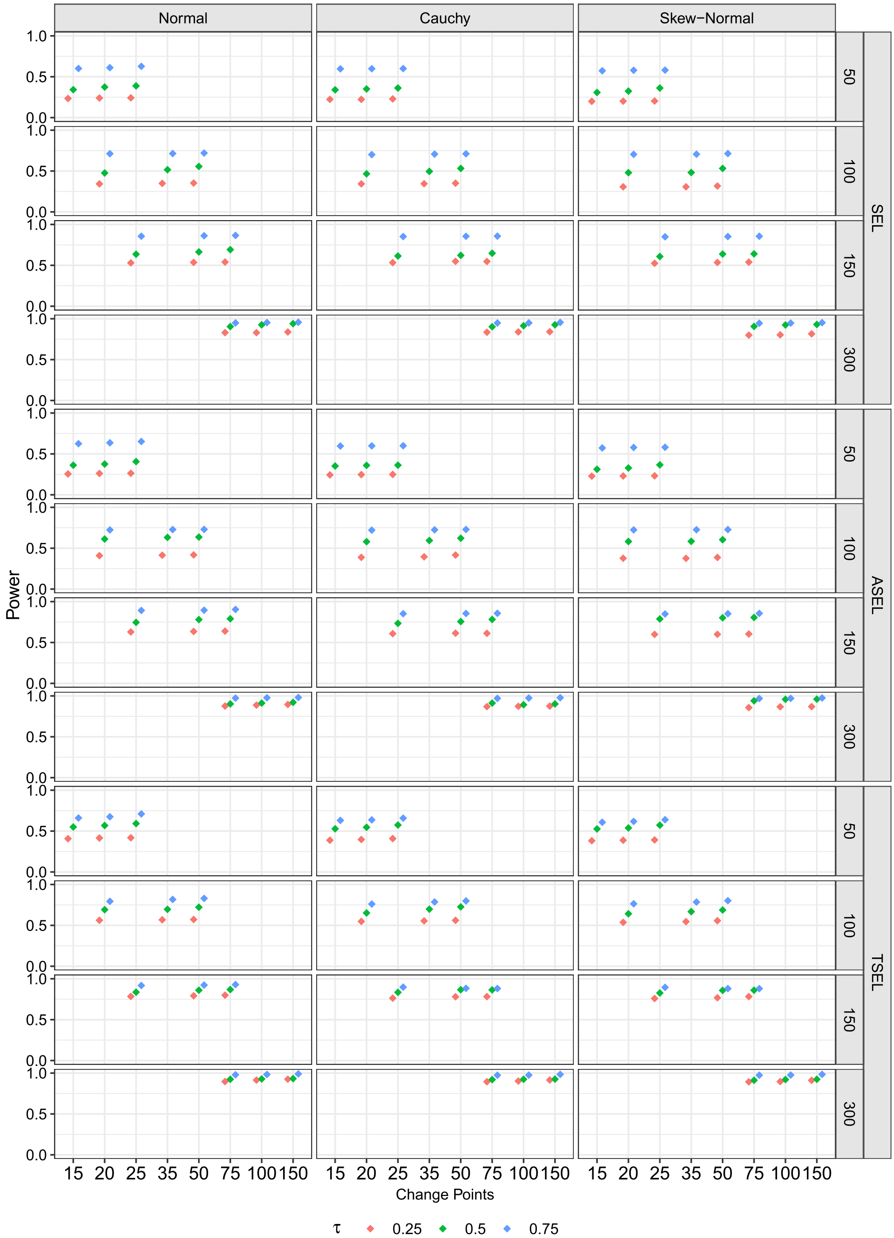

Fig. 2 Power comparison based on SEL, ASEL, and TSEL methods with various change point locations and different sample sizes, $n = 50, 100, 150, 300$

sample size is $n = 50$, the empirical power of the change point location at $k = 30$ would be approximately closer to the empirical power when the change point location is at $k = 20$. Furthermore, we performed a simulation study to assess the impact of the break size, considering different post-change parameter values for β_τ^2, such as 5, 7, 9. We observed that the power tends to increase as the break size increases, even with small sample sizes.

5 Application

In this section, we use the SEL-based quantile approach to detect the structural change in the level of Glycosaminoglycans (GAGs) in *Urine of Children* data set. GAGs are crucial for maintaining cartilage, controlling extracellular signaling, and developing neurons. To identify mucopolysaccharidoses (MPSs), the extracellular concentration of total GAGs has long been employed as a standard measurement. The data set was initially studied by Venables and Ripley (2002). The data set can be obtained from the R package "MASS". The data were collected regarding the amount of a molecule called GAG present in the urine of 314 children, ranging in age from zero to seventeen years. The goal of the study was to create a chart that would assist pediatricians in determining whether a child's GAG concentration is 'normal'. The variables are the age of a child in years and the concentration of GAG. For this analysis, we consider that the outcome variable (y) is GAG and the predictor variable (x) is age. To detect the change point, we first determine the $\tilde{Z}_n$ using the equation (17) with quantile level $\tau = 0.5$. We estimate the change point location at $k = 23$ with a corresponding age of 0.05 years. This indicates that there is a change in the model parameters. Thus, the proposed SEL-based quantile regression method detects a change point with the estimated quantile functions as,

$$Q_{\mathrm{GAG}}(\tau = 0.5|\mathrm{Age}) = \begin{cases} 23.8 + 210 * \mathrm{Age}; & 0 < \mathrm{Age} \leq 0.05 \\ 15.8108197 - 0.8852459 * \mathrm{Age}; & 0.05 < \mathrm{Age} \leq 17.67. \end{cases}$$

Further, the binary segmentation method by Vostrikova (1981) is utilized for any conceivable multiple changes in the data. This method divides the data into homogeneous subsets. However, we found that no additional change points exist. The estimated change point remains the same for quantile levels $\tau = 0.45$ and 0.75. Additionally, we employ the proposed ASEL and TSEL-based methods for quantile regression models with different quantile levels, which also results in the same conclusion (Fig. 3).

6 Discussion

In this paper, we develop an offline EL-based change point detection method to monitor structural changes in QR models. For a fixed change location, we have derived the test statistic which considers the maximum overall possible change locations and it adheres to an asymptotic Extreme value distribution. To remedy the issue that the computational EL problem has no solution, we have adopted the

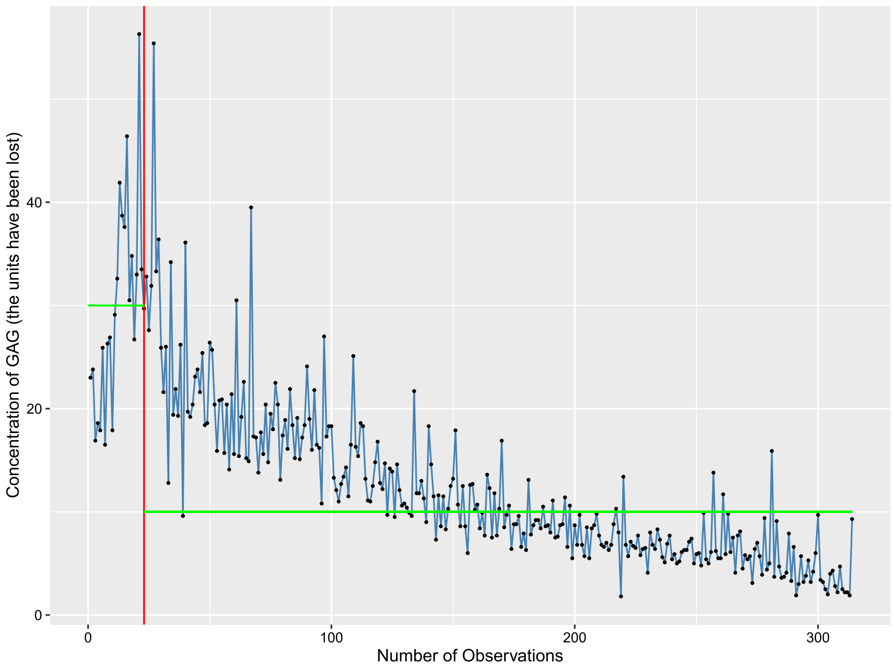

Fig. 3 The level of GAG in Urine of Children data with change point location, $k = 23$

AEL method proposed by Chen et al. (2008) which adds a pseudo term to guarantee the zero vector is within the convex hull, hence assuring that the solution always exists. Furthermore, the TEL method developed by Jing et al. (2017) is used to improve the coverage probabilities in small samples while preserving the properties of the EL method. Since the quantile score function is not differentiable everywhere, we replace it with a smooth function. Thus, the proposed SEL-based change point problem in the QR model is extended using the ASEL and TSEL-based approaches. Monte Carlo simulations have been performed to examine the effectiveness of the three methods including SEL, ASEL, and TSEL with various sample sizes and different change point locations. In the simulation study, we also incorporate different types of error distributions, considering homoskedastic errors, heavy-tailed errors, and skewed errors. Thus, the proposed method can be used to identify changes across the distribution at various quantile levels. Moreover, the proposed method does not require any distributional assumptions of the response or the covariates. The findings show that for the three different distributions, the TSEL approach compares favorably to the other two methods. Finally, a real-data application is provided to illustrate the effectiveness of the proposed change point detection procedure in QR models.

Appendix: Proofs of theorems

Lemma 1 *If H_0 and conditions of Theorem 1 hold, then for all $\zeta > 0$, we can find $C = C(\zeta), T = T(\zeta)$ and $n = n(\zeta)$ such that*

$$P\left(\max_{\frac{T}{n}\le\theta_{nk}\le 1-\frac{T}{n}}\left(\frac{n\theta_{nk}}{\log\log n\theta_{nk}}\right)^{1/2}\left|\left|\frac{\tilde{\lambda}}{\epsilon_k}\right|\right|^{1/2} > C\right) \le \zeta,$$

$$P\left(\max_{\frac{T}{n}\le\theta_{nk}\le 1-\frac{T}{n}}\left(\frac{n\theta_{nk}}{\log\log n\theta_{nk}}\right)^{1/2}||\tilde{\beta}-\beta||_2 > C\right) \le \zeta,$$

$$P\left(n^{-1/2}\max_{\frac{T}{n}\le\theta_{nk}\le 1-\frac{T}{n}} n\theta_{nk}\left|\left|\frac{\tilde{\lambda}}{\epsilon_k}\right|\right|^{1/2} > C\right) \le \zeta,$$

$$P\left(n^{-1/2}\max_{\frac{T}{n}\le\theta_{nk}\le 1-\frac{T}{n}} n\theta_{nk}||\tilde{\beta}-\beta||_2 > C\right) \le \zeta,$$

where $\epsilon_k = \min\{\theta_{nk}, 1-\theta_{nk}\}$.

Proof The proof of this lemma is similar to that of the Lemma 1.2.2 of Csörgö and Horváth (1997) and hence is omitted. □

Lemma 1 gives estimates for the rate of convergence.

Proof Theorem 3.1:

The SEL approach involves the use of a smooth function $\psi_h(\cdot)$ that is differentiable with respect to β_τ. Therefore, Theorem 3.1 can be proven by following similar arguments to those in Liu et al. (2008). According to Lemma A.1 of Liu et al. (2008) and Theorem 2 of Ciuperca and Salloum (2015), it can be seen that our proposed estimators of the parameters are consistent under H_0. To prove the theorem, we first apply Lemma A.3 in Liu et al. (2008) and a justification similar to that in Theorem 1.3.1 of Csörgö and Horváth (1997). The only difference is that, since we require the asymptotic null distribution of $\tilde{Z}_n$, we derive our theorem from Theorem A.3.4 of Csörgö and Horváth (1997).

The overall structure of this proof is similar to those used in linear models but this requires specific modifications due to the SEL method. These modifications primarily involve the smooth function and its differentiability with respect to the parameters. By following the methodologies proposed by Whang (2006) and Wang and Zhu (2011), we can address this issue. Thus, the details are omitted here. □

Proof Theorem 3.2:

First, we are going to show that $-2\log\Lambda_k$ attains its maximum almost surely. For arbitrary small $\frac{\theta_0}{2} > \eta, |\frac{k_0-k}{n}| \ge \eta, -2\log\Lambda_k$ cannot arrive at its maximum with probability approaching 1. Without loss of generality, suppose $k < k_0$ and $\frac{k_0-k}{n} \ge \eta$. Then we have,

$$2\log\Lambda_{k_0} - (-2\log\Lambda_k) = \left(\Lambda_{H_0,k_0} - \Lambda_{H_1,k_0}\right) - \left(\Lambda_{H_0,k} - \Lambda_{H_1,k}\right)$$

Since $\Lambda_{H_1,k} = O_p(1)$, by Theorem 3 of Piyadi Gamage and Ning (2021), for any $\beta^*, \beta \in \mathbb{B}$ and for every fixed parameter $\delta = \beta^* - \beta \neq 0$, there exists a positive constant $c > 0$, we get

$$P\left[\left(-2\log\Lambda_{k_0} - (-2\log\Lambda_k)\right) > cn\right] \longrightarrow 1$$

as $n \longrightarrow \infty$. Since $\Lambda_{H_1,k_0} = O_P(1)$. Because of this, $-2\log\Lambda_k$ cannot reach its maximum with a probability close to 1. Thus, we have $\left|\frac{k_0-\hat{k}}{n}\right| \leq \eta$ with probability approaching to 1, where η is arbitrary. For more details, we refer to the proof of Theorem 3 of Piyadi Gamage and Ning (2021). Then following the Theorem 4 of Ciuperca and Salloum (2015), we can show that $\max_{\theta_{nk}\in\Theta_n^*}\left\{ -2\log\Lambda_k\right\} \longrightarrow \max_{0\leq\theta\leq 1}\omega(\theta)$ almost surely. $-2\log\Lambda_k \longrightarrow \omega(\theta)$, where $\theta = \lim_{n\longrightarrow\infty}\theta_{nk}$. This completes the proof. □

Proof Theorem 3.3:

The ELR test statistic is,

$$-2\log\Lambda_k = \Lambda_{H_0,k_0} - \Lambda_{H_1,k_0}.$$

Under the alternative hypothesis, Λ_{H_1,k_0} follows an asymptotic χ^2 distribution with one degrees of freedom. Thus, $\Lambda_{H_1,k_0} = O_P(1)$. Further, Piyadi Gamage and Ning (2021) showed that, for a positive constant c and under H_1, $P(\Lambda_{H_0,k_0} > cn)\xrightarrow{a.s.} 1$. Thus, the ELR test is consistent. □

Similarly, for ASEL and TSEL methods, following the same approach, we can show that under the same conditions, Theorem 3.2 and 3.3 are valid. Thus, the details are omitted here.

References

Buchinsky M (1998) Recent advances in quantile regression models: a practical guideline for empirical research. J Hum Resour 33(1):88–126

Chen J, Gupta AK (2000) Parametric statistical change point analysis. Birkhäuser, Boston

Chen J, Variyath A, Abraham B (2008) Adjusted empirical likelihood and its properties. J Comput Graph Stat 17(2):426–443

Ciuperca G, Salloum Z (2015) Empirical likelihood test in a posteriori change-point nonlinear model. Metrika 78(8):919–952

Csörgö M, Horváth L (1997) Limit theorems in change-point analysis. Wiley, New York

Gao Q, Zhou X, Feng Y, Du X, Liu X (2021) An empirical likelihood method for quantile regression models with censored data. Metrika 84:75–96

Jing B-Y, Tsao M, Zhou W (2017) Transforming the empirical liklihood towards better accuracy. Can J Stat 45(3):340–352

Koenker R (2004) Quantile regression for longitudinal data. J Mult Anal 91:74–89

Koenker R, Bassett B (1978) Regression quantiles. Econometrica 46:33–50

Koenker R, Geling R (2001) Reappraising medfly longevity: a quantile regression survival analysis. J Am Stat Assoc 96:458–468
Koenker R, Hallock K (2001) Quantile regression. J Econ Perspect 15:143–156
Liu Y, Zou C, Zhang R (2008) Empirical likelihood ratio test for a change-point in linear regression model. Commun Stat Theory Methods 37(16):2551–2563
Owen AB (1988) Empirical likelihood ratio confidence intervals for a single functional. Biometrika 75:237–249
Owen AB (1990) Empirical likelihood confidence regions. Ann Stat 18:90–120
Owen AB (2001) Empirical likelihood. Champan & Hall, New York
Piyadi Gamage RD, Ning W (2020) Inference for long-memory time series models based on modified empirical likelihood. Austrian J Stat 49(5):68–79
Piyadi Gamage RD, Ning W (2021) Empirical likelihood for change point detection in autoregressive models. J Korean Stat Soc 50:69–97
Qin J, Lawless J (1994) Empirical likelihood and general estimating equations. Ann Stat 22:300–325
Tang C, Leng C (2011) Empirical likelihood and quantile regression in longitudinal data analysis. Biometrika 98(4):1001–1006
Tsao M (2013) Extending the empirical likelihood by domain expansion. Can J Stat 41(2):257–274
Venables WN, Ripley BD (2002) Modern applied statistics with S-PLUS, 4th edn. Springer, Berlin
Vostrikova L (1981) Detecting 'disorder' in multidimensional random processes. Sov Math Dokl 24:55–59
Wang H, Zhu Z (2011) Empirical likelihood for quantile regression models with longitudinal data. J Stat Plan Inference 141(4):1603–1615
Wei Y, Pere A, Koenker R, He X (2006) Quantile regression methods for reference growth charts. Stat Med 25(8):1369–1382
Whang Y (2006) Smoothed empirical likelihood methods for quantile regression models. Econom Theory 22:173–205
Yang Y, He X (2012) Bayesian empirical likelihood for quantile regression. Ann Stat 40(2):1102–1131
Zhao Y, Chen F (2008) Empirical likelihood inference for censored median regression model via nonparametric kernel estimation. J Multivar Anal 99:215–231

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.