Annals of the Institute of Statistical Mathematics (2024) 76:333–359
https://doi.org/10.1007/s10463-023-00889-z

Test for conditional quantile change in general conditional heteroscedastic time series models

Sangyeol Lee[1] · **Chang Kyeom Kim**[1]

Received: 8 June 2023 / Revised: 19 October 2023 / Accepted: 24 October 2023 /
Published online: 15 December 2023

Abstract
This study aims to test for detecting a change point in the conditional quantile of general location-scale time series models. This issue is quite important in risk management because the conditional quantile is utilized to measure the value-at-risk or expected shortfall of financial assets. In this paper, we design two types of cumulative sum tests based on the conditional quantiles. Their limiting null distributions are derived under regularity conditions, together with consistency of the proposed tests under the alternative. Monte Carlo simulations demonstrate the good performance of the proposed tests in terms of both stability and power for various time series settings. A real data analysis using the daily returns of the Brent Oil futures also confirms the validity of the tests in real-world applications.

Keywords Change point detection · Conditional heteroscedastic time series models · CUSUM test · Quantile regression · Risk management

1 Introduction

Quantile regression was introduced by Koenker and Bassett (1978) to estimate the conditional quantile functions for regressive models at any given quantiles and has been used to measure the tail risk of the response variable. Later it has been applied to various research fields such as econometrics and finance. We refer to Bloomfield and Steiger (1983), Weiss (1991), Koenker and Zhao (1996), Xiao and Koenker (2009), Lee and Noh (2013), and Noh and Lee (2016). Quantile regression was then extended by Engle and Manganelli (2004) to measure the value-at-risk (VaR) of financial time series using the conditional autoregressive heteroscedastic (ARCH) models. We also refer to Fitzenberger et al. (2013) and Kim and Lee (2016) for modeling the VaR and the expected shortfall (ES) in generalized ARCH (GARCH)-type

✉ Sangyeol Lee
sylee@stats.snu.ac.kr

[1] Department of Statistics, Seoul National University, Seoul 08826, Korea

time series models. In particular, Noh and Lee (2016) particularly considered the quantile regression estimators for the general location-scale time series models, which includes a broad class of models such as ARMA and GARCH models, and derived the consistency and asymptotic normality of such quantile regression estimators.

In this paper, we consider the change point test problem for the conditional quantiles of the general location-scale time series. The change point test for time series, pioneered by the seminal paper of Page (1955), has a long history and has been regarded as one of the influential research topics in time series analysis over past decades. It has been deployed in various fields of research, including economics, finance, statistical process control, and health sciences, where detecting the cause of the structural change is crucial at tracing the phenomenon of status change. We refer to Csörgő and Horváth (1997) and Chen and Gupta (2012) for a general review of the field. In particular, the change point test employing the cumulative sum (CUSUM) method has been acclaimed by a broad range of researchers and practitioners because of its great detection ability, reliability and convenience. We refer to Inclán and Tiao (1994), Kim et al. (2000), Lee et al. (2003), Berkes et al. (2004), Gombay (2008), Oh and Lee (2018), and the papers cited therein, to provide a comprehensive look regarding the development of the CUSUM method. The CUSUM tests have been confirmed to retain their stability and potency to detect changes for GARCH models (Engle 1982 and Bollerslev 1986) as validated in Berkes et al. (2004) and Oh and Lee (2019).

CUSUM tests were originally formulated using the estimators of target parameters of the model, and were devised to compare the discrepancy among sequentially calculated estimates. However, estimate-based CUSUM tests tend to suffer from oversizing, which impairs the overall stability of the test. To circumvent the problem, the residual-based CUSUM test has later been advocated by, e.g., de Pooter and van Dijk (2004), Lee et al. (2004), and Oh and Lee (2019). Residual-based tests regain stability by mitigating the size distortion problem rooted from estimate-based CUSUM tests. Inspired by this fact, we here propose two variants of residual-based CUSUM tests featuring the scheme of quantile regression estimation established in Noh and Lee (2016). Specifically, we propose a test which utilizes the marked empirical process-type CUSUM process, and a CUSUM of squares test based on model residuals.

Change point tests via a CUSUM method, primarily targeting to detect the structural change in conditional quantiles, have been explored in the literature as a form of both retrospective change point test and prospective monitoring scheme, see Qu (2008), Oka and Qu (2011), Zhou et al. (2015), and Ciuperca (2017). Moreover, it has been extended to embrace some time series models, such as the quantile AR model (Su and Xiao 2008). Lee and Kim (2022) recently considered this problem for GARCH(1, 1) models, and proposed two CUSUM tests which perform effectively well detecting changes in tail quantiles in real-world circumstances. Nevertheless, their tests possess some opposite merits and drawbacks. Specifically, it was revealed that one of their tests exhibits a good stability but produces low powers on occasions, and the other test has oversizing problems but produces superior powers. Therefore, our study is not only devoted to the extension of those tests to the general location-scale time series models, but also to the improvement of those two tests through a suitable modification in test statistics and in algorithms for better tunings. Regularity conditions required for establishing the null limiting distributions are also properly tailored to the case of general location-scale time series models. Moreover, we investigate the behavior of our change point tests under the alternative, while imposing milder regularity conditions compared to those of Oh and Lee (2019).

The fundamental difference between the classical tests as in Oh and Lee (2019) and our proposed tests stems from the intrinsic disparity between the parameter estimators and the conditional quantile estimators, although both estimators share similar asymptotic properties (Noh and Lee 2016; Oh and Lee 2019). Indeed, classical tests were constructed using estimators of the individual model parameters, such as the quasi maximum likelihood estimator (QMLE) (Francq and Zakoian 2004). On the other hand, our test is particularly crafted using the conditional quantile estimators, thus is able to adequately harness the information of residuals derived from those estimators. As the approach of quantile regression is semiparametric in principle, the discrepancy particularly prohibits one from simply plugging in the individual parameter estimators into the formula of the conditional quantiles in constructing our test. Indeed, such a heuristic parametric method does not mandatorily guarantee the desired accuracy of the quantile estimators. Also, the quantile regression estimators has robust properties against outliers, which may fail to hold for QMLE. As a whole, the purpose of the quantile change test is conceptually different from the parameter change test, thus it should be properly distinguished.

The rest of this paper is structured as follows. Sections 2 and 3 introduce CUSUM tests for conditional quantile changes in the location-scale time series models, and study their asymptotic properties. Section 4 conducts a simulation study to evaluate the performance of the proposed CUSUM tests. Section 5 analyzes the daily returns of the Brent Oil futures from September 2010 to September 2018 to exhibit its applicability to the real-world financial time series. Finally, Sect. 6 provides the concluding remarks. The proofs are provided in the Appendix.

2 Quantile change point for location-scale models

2.1 Quantile regression estimator

Let us consider a conditional location-scale model of the form:

$$y_t = g_t(\mu_0) + h_t^{1/2}(\theta_0)\eta_t, \quad t \in \mathbb{Z}, \tag{1}$$

where $g_t(\mu) = g(y_{t-1}, y_{t-2}, \dots; \mu)$ is the conditional mean and $h_t(\theta) = h(y_{t-1}, y_{t-2}, \dots; \theta)$ is the conditional variance; $g : \mathbb{R}^\infty \times \Theta_1 \to \mathbb{R}$ and $h : \mathbb{R}^\infty \times \Theta \to \mathbb{R}$ are known measurable functions; $\mu \in \Theta_1 \subset \mathbb{R}^{m_1}$ and $\theta = (\mu^T, \lambda^T)^T \in \Theta = \Theta_1 \times \Theta_2 \subset \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ are unknown parameters; $\theta_0 = (\mu_0^T, \lambda_0^T)^T$ is the true parameter; $\{\eta_t\}$ is a sequence of iid random variables with mean zero and a finite fourth moment. In what follows, we assume that $\{y_t\}$ is strictly stationary and ergodic, η_t is independent of y_s for all $s < t$, and further, $Ey_t = 0$ and $Ey_t^4 < \infty$. For the stationarity and ergodicity, we refer to Bougerol and Picard (1992), Straumann and Mikosch (2006), and Meitz and Saikkonen (2008).

Model (1) includes a broad class of stationary ARMA and GARCH type models. A representative location-scale model is the ARMA(p, q)-GARCH(1,1) model as follows:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \psi_j \epsilon_{t-j} + \epsilon_t,$$
$$\epsilon_t = h_t^{1/2} \eta_t,\ h_t = 1 + \alpha \epsilon_{t-1}^2 + \beta h_{t-1}, \tag{2}$$

where $\phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i \neq 0$ and $\psi(z) = 1 + \sum_{j=1}^{q} \psi_j z^j \neq 0$ for all $|z| \leq 1$ in the complex plane, and η_t's are iid with mean zero and variance $\omega > 0$. Here, we assume $\alpha, \beta \geq 0$ and $\alpha\omega + \beta < 1$, which ensure the stationarity (causality) and invertibility; see $(A1)'$ in Noh and Lee (2016).

Below, we present the scheme of the quantile regression estimator for the strictly stationary location-scale model in (1), introduced in Noh and Lee (2016). For $0 < \tau < 1$, the τth quantile of y_t conditional on the past observations is given by

$$q_t(\vartheta_0) = g_t(\mu_0) + \xi_0 h_t^{1/2}(\theta_0), \tag{3}$$

where $\xi_0 = F_\eta^{-1}(\tau)$ is assumed to be nonzero for model identifiability, with $F_\eta(\cdot)$ being the distribution function of η_1. We denote by $\vartheta_0 = (\xi_0, \theta_0^T)^T$ the true parameter vector. For $\vartheta \in \Theta^*$, where Θ^* is a subset of $\mathbb{R} \times \Theta$, we express

$$q_t(\vartheta) = g(y_{t-1}, y_{t-2}, \dots; \mu) + \xi h^{1/2}(y_{t-1}, y_{t-2}, \dots; \theta), \tag{4}$$

which is approximated with

$$\begin{aligned}\tilde{q}_t(\vartheta) &:= \tilde{g}_t(\mu) + \xi \tilde{h}_t^{1/2}(\theta) \\ &= g(y_{t-1}, \dots, y_1, c_1, c_1, \dots; \mu) + \xi h^{1/2}(y_{t-1}, \dots, y_1, c_2, c_2, \dots; \theta),\end{aligned}$$

wherein all y_t, $t \leq 0$, are set to constants c_1 and c_2 (usually 0). Then, in light of (3), the τth quantile regression estimator for model (1) is defined by

$$\hat{\vartheta}_n = \underset{\vartheta \in \Theta^*}{\operatorname{argmin}} \frac{1}{n} \sum_{t=1}^{n} \rho_\tau(y_t - \tilde{q}_t(\vartheta)), \tag{5}$$

where $\rho_\tau(u) = u(\tau - I(u \leq 0))$. Noh and Lee (2016) established that $\hat{\vartheta}_n$ is a consistent estimator of ϑ_0 and is asymptotically normal as $n \to \infty$ under regularity conditions. The following conditions are their reformulations and implicate the conditions in Theorems 1 and 2 of Noh and Lee (2016) except for their *(N4)*(iii); see Remark 2.

(A1) Θ^* is compact and ϑ_0 is an interior point of Θ^*.

(A2) $h_t(\theta) \geq c$ uniformly in t and θ for some positive constant c.

(A3) If $q_t(\vartheta) = q_t(\vartheta_0)$ a.s. for some t and $\vartheta \in \Theta^*$, then $\vartheta = \vartheta_0$.

(A4) The distribution function F_η of η_1 has a positive continuous density f_η and $\sup_x f_\eta(x) < \infty$.

(A5)
$g_t(\mu)$ and $h_t(\theta)$ are twice continuously differentiable on Θ and the following moment conditions hold:

(i) $E\big(\sup_{\theta\in\theta} |g_t(\mu)|\big)^4 < \infty$, $E\big(\sup_{\theta\in\Theta} \|\frac{\partial g_t(\mu)}{\partial \mu}\|\big)^4 < \infty$;

(ii) $E\big(\sup_{\theta\in\Theta} |h_t(\theta)|\big)^2 < \infty$, $E\big(\sup_{\theta\in\Theta} \|\frac{1}{h_t(\theta)}\frac{\partial h_t(\theta)}{\partial \theta}\|\big)^4 < \infty$;

(iii) $E\big(\sup_{\theta\in\Theta} \|\frac{\partial^2 g_t(\mu)}{\partial\mu\partial\mu^T}\| + \sup_{\theta\in\Theta} \|\frac{1}{h_t(\theta)}\frac{\partial^2 h_t(\theta)}{\partial\theta\partial\theta^T}\|\big)^2 < \infty$.

(A6) Both g and h and their first derivatives belong to the class $\mathcal{A}$. Herein, $f : \mathbb{R}^\infty \times \Theta \to \mathbb{R}$ is said to belong to the class $\mathcal{A}$ if for all $t \geq 1$,

$$\sup_{\theta\in\Theta} \left\| f(y_t, y_{t-1}, \dots ;\theta) - f(y_t, \dots, y_1, c, c, \dots ;\theta) \right\| \leq V\rho^t \quad \text{a.s.}$$

for a constant c, where V is an integrable random variable and $\rho \in (0, 1)$ is a constant.

(A7) The following matrix $J(\tau)$ is positive definite:

$$J(\tau) = E\left[\frac{1}{h_t(\theta_0)} \frac{\partial q_t(\vartheta_0)}{\partial\vartheta} \frac{\partial q_t(\vartheta_0)}{\partial\vartheta^T}\right]. \tag{6}$$

Remark 1 The above assumptions with (A3) replaced by $(A3)'$: $g_t(\mu) = g_t(\mu_0)$ and $h_t(\theta) = h_t(\theta_0)$ a.s. for some t if and only if $\theta = \theta_0$, can yield the consistency and asymptotic normality of the Gaussian QMLE of θ_0 as verified by Oh and Lee (2019), for which, however, (A4) becomes redundant. In fact, under certain circumstances, (A3) and $(A3)'$ are equivalent. Particularly, (A5) implies $E\Big(\sup_{\vartheta\in\Theta^*} \|\frac{\partial q_t(\vartheta)}{\partial\vartheta}\|^2\Big) + E\Big(\sup_{\vartheta\in\Theta^*} \|\frac{\partial^2 q_t(\vartheta)}{\partial\vartheta\partial\vartheta^T}\|\Big) < \infty$. Stationary ARMA-GARCH and pure asymmetric GARCH (AGARCH) (or GJR-GARCH) models satisfy (A1)–(A7) under fairly mild conditions; see Sect. 3 of Noh and Lee (2016).

More specifically, the asymptotic results of Noh and Lee (2016) are summarized as follows, which will be utilized for analyzing the asymptotic behavior of the CUSUM test in the next section.

Proposition 1 *Under* (A1)–(A7), $\hat{\vartheta}_n$ *converges to* ϑ_0 *a.s. and*

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta_0) \xrightarrow{d} N\left(0, \frac{\tau(1-\tau)}{f_\eta^2(F_\eta^{-1}(\tau))} J(\tau)^{-1} V(\tau) J(\tau)^{-1}\right),$$

where $J(\tau)$ is given in (6) and

$$V(\tau) = E\left[\frac{\partial q_t(\vartheta_0)}{\partial \vartheta} \frac{\partial q_t(\vartheta_0)}{\partial \vartheta^T}\right].$$

Remark 2 Noh and Lee (2016) assumed in their condition (N4)(iii) that the second derivatives of $g_t(\mu)$ and $h_t(\theta)$ belong to $\mathcal{A}$. However, this condition is redundant since the approximating assumption regarding the first derivatives of $g_t(\mu)$ and $h_t(\theta)$ is sufficient for the verification. This can be easily checked by looking over the proof of Lemma A.3 of Noh and Lee (2016).

2.2 CUSUM test

In this subsection, we consider the problem of testing a change in the conditional quantile $q_t(\vartheta_0)$. As the conditional quantile formula depends on ϑ, we convert this problem into the change point testing problem for ϑ. Henceforth, we set up the following hypotheses:

$$H_0 : \vartheta \text{ is constant as } \vartheta_0 \text{ over } t = 1, \dots, n. \quad \text{vs.} \quad H_1 : \text{ not } H_0. \tag{7}$$

However, when testing for H_0 vs. H_1, rather than using the classical parameter change test of Oh and Lee (2019), we employ another type of CUSUM test that reflects the features of conditional quantiles as follows:

$$\begin{aligned} \hat{T}_n &:= \max_{1\le k\le n} \hat{T}_n(k) \\ &= \frac{1}{\sqrt{n}\hat{\tau}_n} \max_{1\le k\le n} \left| \sum_{t=1}^{k} |y_{t-1}|^\gamma \psi_\tau(y_t - \hat{q}_t) - \frac{k}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma \psi_\tau(y_t - \hat{q}_t) \right| \end{aligned} \tag{8}$$

with $0 \le \gamma \le 2$, $\psi_\tau(u) = \tau - I(u \le 0)$, $\hat{q}_t = \tilde{g}_t(\hat{\mu}_n) + \hat{\xi}_n \tilde{h}_t^{1/2}(\hat{\theta}_n)$, and

$$\hat{\tau}_n^2 = \frac{1}{n} \sum_{t=1}^{n} y_{t-1}^{2\gamma} \left(\psi_\tau(y_t - \hat{q}_t)\right)^2 - \left(\frac{1}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma (\psi_\tau(y_t - \hat{q}_t))\right)^2,$$

which is similar to the test proposed in Zhou et al. (2015), Ciuperca (2017), and Lee and Kim (2022). It is notable that unlike Lee and Kim (2022), who used $\gamma = 2$, we allow to use any $\gamma \in [0, 2]$ not only to enhance the flexibility of the test, but also to significantly augmenting the detection ability of the test in many circumstances. Moreover, our simulation study suggests that a small γ is favored for better detecting scale parameter changes, while a large γ is favored for handling location parameter changes.

Under (A1)–(A7) and certain additional assumptions, the limiting null distribution of the test can be shown as follows (see the Appendix for the proof): As $n \to \infty$,

$$\text{(i)} \quad \hat{T}_n - T_n = o_P(1), \tag{9}$$

where

$$T_n = \frac{1}{\sqrt{n}\tau} \max_{1\le k\le n} \left| \sum_{t=1}^{k} |y_{t-1}|^\gamma \psi_\tau(y_t - q_t(\vartheta_0)) - \frac{k}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma \psi_\tau(y_t - q_t(\vartheta_0)) \right|$$

with $\tau^2 = Var\big(|y_1|^\gamma \psi_\tau(y_2 - q_2(\vartheta_0))\big)$, and

$$\text{(ii)} \quad T_n \xrightarrow{d} \sup_{0\le s\le 1} |B^\circ(s)|,$$

where B° is a Brownian bridge, due to Donsker's invariance principle (Billingsley 1968).

Theorem 1 *Suppose that* (A1)–(A7) *hold. Moreover, assume* $\sup_x |f_\eta'(x)| < \infty$.

Then, under H_0, *as* $n \to \infty$,

$$\hat{T}_n \xrightarrow{d} \sup_{0\le s\le 1} |B^\circ(s)|.$$

Remark 3 Lee and Kim (2022) assumed a more stringent condition that $E|y_1|^{4+\delta} < \infty$ for some $\delta > 0$ in GARCH(1, 1) models, which is no longer necessary in our analysis. The theorem can be proven through manipulating a martingale sequence similarly to the one for the GARCH(1, 1) model. However, the approach herein is more sophisticated than that of Lee and Kim (2022) in that the martingale sequence is equipped with a stopping rule scheme to overcome the restriction of the 4th moment condition. This part is of independent interest.

Aside form $\hat{T}_n$, we can employ another CUSUM test based on $\tilde{g}_t(\hat{\mu}_n)$ and the squared residuals $\hat{\eta}_t{}^2$ with

$$\hat{\eta}_t = (y_t - \tilde{g}_t(\hat{\mu}_n))/\tilde{h}_t^{1/2}(\hat{\theta}_n).$$

Specifically, we utilize the following test statistics (Lee 2020):

$$\tilde{T}_n = \max_{1\le k\le n} \left\{ \tilde{T}_{n,1}(k) + \tilde{T}_{n,2}(k) \right\},$$

where

$$\begin{aligned} \tilde{T}_{n,1}(k) &= \frac{1}{n\hat{\tau}_{1,n}^2} \left| \sum_{t=1}^{k} \tilde{g}_t(\hat{\mu}_n)\hat{\eta}_t - \left(\frac{k}{n}\right) \sum_{t=1}^{n} \tilde{g}_t(\hat{\mu}_n)\hat{\eta}_t \right|^2, \\ \tilde{T}_{n,2}(k) &= \frac{1}{n\hat{\tau}_{2,n}^2} \left| \sum_{t=1}^{k} \hat{\eta}_t^2 - \left(\frac{k}{n}\right) \sum_{t=1}^{n} \hat{\eta}_t^2 \right|^2 \end{aligned} \tag{10}$$

with

$$\hat{\tau}_{1,n}^2 = \frac{1}{n}\sum_{t=1}^{n}\tilde{g}_t(\hat{\mu}_n)^2\hat{\eta}_t^2 - \left(\frac{1}{n}\sum_{t=1}^{n}\tilde{g}_t(\hat{\mu}_n)\hat{\eta}_t\right)^2,$$

$$\hat{\tau}_{2,n}^2 = \frac{1}{n}\sum_{t=1}^{n}\hat{\eta}_t^4 - \left(\frac{1}{n}\sum_{t=1}^{n}\hat{\eta}_t^2\right)^2.$$

The primary reason why we consider $\tilde{T}_n$ is that it substantially augments the detection ability regarding the change in location parameters compared to $\hat{T}_n$, as further revealed in the simulation study in Sect. 4. The following result shows the weak convergence of the residual-based CUSUM of squares test to the supremum of a Brownian bridge, which can be verified similarly to Oh and Lee (2019) and Lee (2020) and the proof is omitted for brevity.

Theorem 2 *Suppose that* (*A*1)–(*A*7) *hold. Then, under* H_0, *as* $n \to \infty$,

$$\tilde{T}_n \xrightarrow{d} \sup_{0\le s\le 1}\|B_2^\circ(s)\|^2,$$

where B_2° *denotes a two-dimensional Brownian bridge.*

In circumstances where the underlying model is a pure GARCH-type, namely, $g_t(\mu) \equiv 0$ for all $t \geq 1$, we alternatively consider the test statistics only consisting of $\tilde{T}_{n,2}(k)$ to particularly examine the change in scale parameters. Specifically, we consider

$$\tilde{T}_{n,2} := \max_{1\le k\le n}\tilde{T}_{n,2}(k).$$

Corollary 1 *Suppose that* (*A*1)–(*A*7) *hold. Then, under* H_0, *as* $n \to \infty$,

$$\tilde{T}_{n,2} \xrightarrow{d} \sup_{0\le s\le 1}|B^\circ(s)|^2.$$

Although $\tilde{T}_n$ and $\tilde{T}_{n,2}$ can precisely detect parameter changes in general, our unreported simulation study indicates that they can occasionally suffer from mild size distortions. As a remedy, we employ a test that uses $\check{T}_n$ and $\check{T}_{n,2}$ below that slightly modifies $\hat{\tau}_{2,n}^2$ to facilitate the test to be much more stable, and report these results in Sect. 4. The following shows that the Brownian bridge result is valid for $\check{T}_n$, the proof of which is omitted for brevity.

Theorem 3 *Suppose that* (*A*1)–(*A*7) *hold. Then, as* $n \to \infty$,

$$\check{T}_n := \max_{1\le k\le n}\left\{\tilde{T}_{n,1}(k) + \check{T}_{n,2}(k)\right\} \xrightarrow{d} \sup_{0\le s\le 1}\|B_2^\circ(s)\|^2$$

and

$$\check{T}_{n,2} := \max_{1\le k\le n} \check{T}_{n,2}(k) \xrightarrow{d} \sup_{0\le s\le 1} |B^\circ(s)|^2,$$

where $\tilde{T}_{n,1}(k)$ *is defined in* (10) *and* $\check{T}_{n,2}(k)$ *is defined as*

$$\check{T}_{n,2}(k) = \frac{1}{n\check{\tau}^2_{2,n}} \left| \sum_{t=1}^{k} \hat{\eta}_t^2 - \left(\frac{k}{n}\right) \sum_{t=1}^{n} \hat{\eta}_t^2 \right|^2.$$

Here, $\check{\tau}^2_{n,2} = \max(\dot{\tau}_n, \nu_n)$, ν_n *is a sequence of positive real numbers decaying to* 0, *and* $\dot{\tau}_n = \hat{\gamma}_n(0) + 2\hat{\gamma}_n(1)$ *with*

$$\hat{\gamma}_n(j) = \frac{1}{n} \sum_{t=1}^{n-j} \left(\hat{\eta}^2_{t+j} - \frac{1}{n}\sum_{t=1}^{n} \hat{\eta}_t^2\right)\left(\hat{\eta}_t^2 - \frac{1}{n}\sum_{t=1}^{n}\hat{\eta}_t^2\right), \quad j = 0, 1.$$

In practice, κ_n can be negligibly small, as it does not influence the overall performance of $\check{T}_n$ in terms of both size and power. For instance, κ_n can be a sequence that decays rapidly toward zero (e.g., $\kappa_n = 0.001n^{-2}$).

3 Consistency of CUSUM test

In this section, we investigate the consistency of the proposed tests and substantiate it under regularity conditions via using the framework in Sect. 2.2 of Oh and Lee (2019). For a relevant paper, see also Kirch and Kamgaing (2012). In this scenario, the null and alternative hypotheses are set up as follows:

$$\mathcal{H}_0 : \vartheta \text{ is constant as } \vartheta_1 \text{ over } t = 1, \dots, n \text{ vs.}$$
$$\mathcal{H}_1 : \vartheta \text{ is constant as } \vartheta_1 \text{ over } t = 1, \dots, k_0 \text{ and changes to } \vartheta_2 \text{ for } t = k_0 + 1, \dots, n,$$

where $k_0 = [n\kappa_0]$ with $\kappa_0 \in (0, 1)$ and $\vartheta_1 \ne \vartheta_2$. Namely, under $\mathcal{H}_1$, $\{y_1, \dots, y_{k_0}\}$ follows Model (1) with $\vartheta = \vartheta_1$ and $\{y_{k_0+1}, \dots, y_n\}$ follows Model (1) with $\vartheta = \vartheta_2$. Specifically, we denote by $y_t = y_{t1}$ for $t = 1, \dots k_0$ and by $y_t = y_{t2}$ for $t = k_0 + 1, \dots, n$. All the symbols g_{ti}, h_{ti}, q_{ti}, $i = 1, 2$ (for example, $q_{ti}(\vartheta) = g_{ti}(\mu) + \xi h_{ti}^{1/2}(\theta)$) with $g_{ti}(\mu) = g(y_{ti}, y_{t-1,i}, \dots; \mu)$ and $h_{ti}(\theta) = h(y_{ti}, y_{t-1,i}, \dots; \theta)$), and their variants are accordingly defined, and the regularity conditions in (A1)–(A7) are assumed to hold for both $\{y_{t1}\}$ and $\{y_{t2}\}$. Specifically, (A5) is modified for $\{y_{t1}\}$ as follows: for $t \le k_0$,

$$\sup_{\theta\in\Theta} ||f(y_{t1}, \dots, y_{11}, c, c \dots; \theta) - f(y_{t1}, y_{t-1,1}, \dots; \theta)|| \le \rho^t V \text{ a.s.}$$

Also, we assume that g, h and their first derivatives belong to the class $\mathcal{A}'$, where $f \in \mathcal{A}'$ implicates that for $t > k_0$,

$$\sup_{\theta \in \Theta} || f(y_{t2}, \ldots, y_{k_0+1,2}, y_{k_0,1}, \ldots, y_{11}, c, c \ldots ; \theta) - f(y_{t2}, y_{t-1,2}, \ldots ; \theta) || \leq \rho^{t-k_0} V \text{ a.s.} \quad (11)$$

for constant c. This condition will be harnessed for investigating the asymptotic properties regarding $\tilde{q}_{t2}$ and $\hat{q}_{t2}$, as seen in the proofs of Theorems 2 and 3 of Oh and Lee (2019). Moreover, like in their study, the following conditions are imposed:

(B1) Under $\mathcal{H}_1$, there exists $\vartheta_0^* = (\xi_0^*, \theta_0^{*T})^T \in \Theta^*$ such that $\hat{\vartheta}_n = \vartheta_0^* + o_P(1)$, where $\hat{\vartheta}_n$ is the one in (5).

(B2) Let $d_i = E|y_{t-1,i}|^\gamma \psi_\tau(y_{ti} - q_{ti}(\vartheta_0^*))$, then $d_1 \neq d_2$.

These conditions result in the following, the proof of which is provided in the Appendix.

Theorem 4 *Let $\hat{T}_n$ be the one in (8) and $\hat{k}_n = \operatorname{argmax}_{1 \leq k \leq n} \hat{T}_n(k)$, and assume that (B1) and (B2) hold. Then, under $\mathcal{H}_1$, we have $\hat{T}_n \to \infty$ and $\hat{k}_n / n \to \kappa_0$ in probability as $n \to \infty$.*

Remark 4 It can be checked from (5) that (B1) holds if $\kappa_0 E\{\rho_\tau(y_{t1} - q_{t1}(\vartheta))\} + (1 - \kappa_0) E\{\rho_\tau(y_{t2} - q_{t2}(\vartheta))\}$ for $\vartheta \in \Theta^*$ has its maximum uniquely at an interior point ϑ_0^* of Θ^*, see Appendix. A similar condition was also assumed in Lemma 5 of Oh and Lee (2019) as the form of the assumption (C1)(i) and (C2)(i) of their study for investigating the consistency of the CUSUM test for location-scale models, as analytically verifying the presence of such ϑ_0^* or finding a suitable sufficient condition could be in general quite difficult except for some basic models like AR models. Therein, a more stringent condition like the $\sqrt{n}$-consistency of $\hat{\vartheta}_n$ to ϑ_0^* was assumed, which, however, can be relaxed to the one in (B1).

To deal with the consistency of $\tilde{T}_n$, we impose the conditions (B1) and

(B3) Let $d_i^{(1)} = E g_{ti}(\mu_0^*) \eta_{ti}(\theta_0^*)$ and $d_i^{(2)} = E \eta_{ti}^2(\theta_0^*)$, where $\eta_{ti}(\theta) = (y_{ti} - g_{ti}(\mu)) / h_{ti}^{1/2}(\theta)$. Then, at least, one of the two conditions $d_1^{(1)} \neq d_2^{(1)}$ and $d_2^{(2)} \neq d_2^{(2)}$ holds true.

These conditions ensure the consistency of $\tilde{T}_n$, the proof of which is omitted as it is straightforward and essentially follows the same lines as those in Oh and Lee (2019). In particular, the consistency of $\tilde{T}_{n,2}$ holds if $d_1^{(2)} \neq d_2^{(2)}$. The case of $\check{T}_n$ can be similarly handled.

4 Simulation study

This section evaluates the performance of the proposed change point tests $\hat{T}_n$, $\tilde{T}_n$ and $\check{T}_n$ for various location-scale models. Section 4.1 describes the experiment settings and evaluate the overall quality of the quantile estimates presented in (5). Section s4.2 summarizes the performance measured in terms of the empirical sizes and powers and reports the accuracy regarding the location of the change point.

4.1 Preliminaries

We evaluate the performance of the proposed change point test under various settings of location-scale time series models. For this task, we employ the ARMA(1, 1)-GARCH(1, 1) model in (2) and the GJR-GARCH(1, 1) model as follows:

$$y_t = h_t^{1/2}\eta_t, \quad h_t = 1 + \alpha_1 {y_{t-1}^{+}}^2 + \alpha_2 {y_{t-1}^{-}}^2 + \beta_1 h_{t-1},$$
$$y_t^{+} = \max(y_t,\ 0), \quad y_t^{-} = -\min(y_t,\ 0), \quad \alpha_1 \geq 0, \quad \alpha_2 \geq 0, \quad \beta \geq 0$$

[cf. Noh and Lee (2016)]. Herein, η_t is generated from a normal distribution with variance ω, implying that the quantile parameter ξ is approximately −1.282 when $\omega = 1$ and $\tau = 0.1$, and we consider $\tau \in \{0.9, 0.1\}$. For the GJR-GARCH model, we consider the three cases:

- *Case 1*. $(\omega, \alpha_1, \alpha_2, \beta) = (0.3, 0.5, 0.1, 0.3)$;
- *Case 2*. $(\omega, \alpha_1, \alpha_2, \beta) = (0.2, 0.3, 0.7, 0.3)$;
- *Case 3*. $(\omega, \alpha_1, \alpha_2, \beta) = (0.7, 0.1, 0.5, 0.6)$.

Also, for the ARMA-GARCH model, we consider the following:

- *Case 4*. $(\phi, \psi, \omega, \alpha, \beta) = (0.2, 0.2, 0.3, 0.3, 0.3)$;
- *Case 5*. $(\phi, \psi, \omega, \alpha, \beta) = (-0.5, -0.2, 0.5, 0.2, 0.3)$;
- *Case 6*. $(\phi, \psi, \omega, \alpha, \beta) = (-0.3, 0.6, 0.6, 0.2, 0.7)$.

Notice that the former three cases are variants of the pure GARCH-type model, namely, $g_t(\mu) \equiv 0$. Therefore, we employ $\tilde{T}_{n,2}$ and $\check{T}_{n,2}$ rather than $\tilde{T}_n$ and $\check{T}_n$ when conducting the experiment.

Prior to the analysis of the proposed change point tests, we first assess the performance of the conditional quantile estimator in (5), as the quality of the estimators is critical in obtaining well-behaved test statistics. For this experiment, we independently generate 1,000 time series of length $2n$ with $n = 500,\ 1,000$ and $1,500$, then evaluate the bias, sample standard deviation (SD), and root mean-squared error (RMSE) of the estimators $(\alpha_1, \alpha_2, \beta, \xi)$ for Cases 1 and 2, and $(\phi, \psi, \alpha, \beta, \xi)$ for the latter two cases, analogous to Lee and Noh (2013). Note that for all cases with $\tau = 0.9$, the true quantile parameter ξ is approximately 0.256, 0.384 and 0.641 for $\omega = 0.2,\ 0.3$ and 0.5, respectively.

To further evaluate the quality of the quantile estimates, we additionally present the results of two backtesting measures, namely, Kupiec's unconditional coverage test

(Kupiec 1995) and Christoffersen's conditional coverage test (Christoffersen 1998). To perform these tests, setting

$$I_t = I(\hat{q}_t(\tau) < \tau), \quad 1 \le t \le 2n,$$

for $\tau \in (0, 1)$, where $I(\mathbf{A})$ is the indicator function for an event $\mathbf{A}$, we construct a likelihood ratio test using I_t and evaluate whether $\{I_t\}$ can be considered as iid random variables from a Bernoulli(τ) distribution. We refer to Campbell (2006) and Lee and Noh (2010).

Tables S.1 and S.2 report the bias, standard deviation (SD), and RMSE of the quantile estimates, and Table S.3 summarizes the results of two backtesting methods mentioned above. These tables all are available in the supplementary material. Notice that these results indicate that some parameters, especially, those in Cases 3 and 6, are somewhat biased and are shown to converge more slowly than the other parameters, which are analogous to the results in Noh and Lee (2016). We speculate that this estimation bias is the byproduct of two factors: (i) the parameter β of both Cases 3 and 6 is relatively large, resulting in the model being nearly nonstationary compared to the other cases, and (ii) the objective function in (5) is not convex, thereby rendering the parameters to converge toward the local minimum. Nevertheless, these results still substantiate that most estimators are ultimately consistent, as both the bias and variance diminishes as n tends to ∞. Moreover, both backtesting measures illuminate that the quality of predicted quantiles $\hat{q}_t(\tau)$ are not undermined for both upper and lower quantiles. These results strongly advocate the validity of the quantile regression estimator of (5) in constructing the change point tests.

4.2 Experiment results

We here report the empirical sizes and powers of the change point tests introduced in Sect. 2.2 under various circumstances. The procedure of the simulation is given as follows. Analogously to the preceding analysis, we first generate time series of length $2n$, and then consider the case that the change point is located at the center of the time series. When computing empirical sizes and powers, we obtain the residuals by fitting the quantile regression model in (5), then, respectively, compute $\hat{T}_n$, $\tilde{T}_{n,2}$ and $\check{T}_{n,2}$ for Cases 1–3, and also $\hat{T}_n$, $\tilde{T}_n$ and $\check{T}_n$ for Cases 4–6. Subsequently, at the significance level of $\alpha_s = 0.05$, we reject the null hypothesis of (7) if the values of the tests are greater than 1.3397 for $\hat{T}_n$, $\tilde{T}_{n,2}$ and $\check{T}_{n,2}$, and also 2.495 for $\tilde{T}_n$ and $\check{T}_n$. These critical values are obtained through Monte Carlo simulations, and this procedure is iterated 1000 times to obtain the empirical sizes and powers.

Using $n = 500$, 1000 and 1500, we inspect the empirical power when the parameters or the innovation distribution for η_t changes to a normal mixture as follows:

$$\eta_t = Z_{k,l}(\omega) := 0.9X + 0.1Y, \ X \sim N(0, \omega), \ Y \sim N(k, l^2\omega).$$

To illustrate, the distribution of η_t changing from $N(0, \omega)$ to that of $Z_{0,7}(\omega)$ when $\omega = 0.3$ actually means the parameter ξ changing from approximately 0.384 to 0.615 when $\tau = 0.9$. Despite the change in any model parameters can contribute to the change in conditional quantiles, the change of η_t particularly reflects the circumstance where the change in conditional quantiles is solely induced from the change of ξ, not from the model parameters. Here, we consider $(k, l) = (0, 5), (0, 7)$ for the GJR-GARCH model, and $(k, l) = (0, 4.5)$ for the ARMA-GARCH model. Moreover, for the ARMA-GARCH model, we additionally consider $(k, l) = (1, 3)$ and $(-1, 3)$ for the upper and lower quantiles, respectively. Throughout the subsection, we abbreviate $Z_{k,l}(\omega)$ as $Z_{k,l}$ and the zero-mean variant $Z_{0,l}(\omega)$ as Z_l for simplicity. Moreover, all the tables that describe the results can be found in the supplementary material.

Tables S.4 to S.6, respectively, portray the empirical sizes and powers for the GJR-GARCH model for $\tau \in \{0.9, 0.1\}$, namely, Cases 1–3. The size in general is stabilized around the prescribed nominal level α_s for all tests, which validates that all three tests can make a stable test even when the sample size is relatively small. When comparing empirical powers of $\hat{T}_n$, the results of $\gamma = 0.2$ and 0.5 appear to surpass those of $\gamma = 1$ and 2 by a large margin and also to be comparable with those of $\tilde{T}_{n,2}$ and $\check{T}_{n,2}$ in many circumstances. This finding indicates that $\gamma = 2$ may dilute information that ψ_τ possesses, and we can massively augment the detection ability of $\hat{T}_n$ by adopting a small γ, which is indeed a substantial improvement over the results of Lee and Kim (2022). However, both $\tilde{T}_{n,2}$ and $\check{T}_{n,2}$ still exhibit their superiority over $\hat{T}_n$ in cases other than ω experiencing a change, including the case of the quantile parameter ξ. In short, change point tests using $\tilde{T}_{n,2}$, $\check{T}_{n,2}$ and $\hat{T}_n$ with small γ are generally preferable when one primarily intends to detect a change in the conditional volatility of tail quantiles.

Tables S.7 to S.9 summarize the results of the ARMA-GARCH model, namely, Cases 4–6. To elaborate, analogously to the GJR-GARCH model case, the tests appear to retain their stability in these cases, aside from a mild downwards size distortion, which can be witnessed when employing $\tilde{T}_n$. Also, for $\hat{T}_n$, it is seen that adopting different values of γ yields distinctive results. Specifically, small values of γ are favored for detecting a change in scale or quantile parameters ω, α, β and ξ. In contrast, large values of γ, especially those larger than 1, are more effective in detecting a change in location parameters, although they seem to be underpowered under some circumstances. Moreover, as $\tilde{T}_n$ and $\check{T}_n$ are particularly augmented to detect a location parameter change as well, they produce stellar performance across all parameter changes.

Most notably, Tables S.6 and S.9 exhibit that the size of $\hat{T}_n$-based tests, namely, variants of $\hat{T}_n$, is generally much more stable than $\tilde{T}_{n,2}$ and $\check{T}_{n,2}$. For instance, the latter two tests suffer from upwards size distortions up to 0.1 in case n is small, whereas the size of $\hat{T}_n$-based tests stabilize around the nominal level of 0.05. As illustrated in the aforementioned tables, this phenomenon is commonly observed when the parameter β associated with GARCH models becomes moderately large, which can often be observed in real-world financial time series. This finding consolidates that both tests have a complementary relationship and so is recommended to be implemented simultaneously in practice, as the former can be utilized to deter false alarms, while the latter is useful when detecting subtle structural changes.

Finally, Tables S.10 and S.11 report the mean and SD for the location of the change point in Cases 1 and 3 when $n = 1000$, which indicates that the change point is located at $t = 1001$. Here, we specifically utilize $\hat{T}_n$ with $\gamma = 2$ for location parameter changes, and use $\gamma = 0.2$ for scale and quantile parameter changes. Although all tests well capture the location of a change, $\tilde{T}_n$ and $\check{T}_n$ particularly have a smaller variance across all instances compared to $\hat{T}_n$. This indicates $\tilde{T}_n$ and $\check{T}_n$ can be more useful when precisely pinpointing the exact period of the structural change in practice.

The results unravel two distinct advantages of our change point tests: (i) the tests respond well to small and subtle changes of the underlying structure; (ii) they can effectively detect the decrease in volatility of tail quantiles. In Table S.5, empirical powers tend to one when n becomes large even when the degree of the parameter change is small, which is well reflected in cases that ω changes from 0.2 to 0.25 or 0.35, or β changes from 0.3 to 0.05. Remarkably, $\tilde{T}_n$ and $\check{T}_n$ are shown to produce high powers even when the length of observations is relatively short, namely $n = 500$. In addition, the power rapidly approaches one when the overall volatility diminishes, which is well illustrated when ω decreases from 0.2 to 0.15 or 0.1. In a nutshell, the two discoveries further solidify the validity of our change point tests.

In addition, $\hat{T}_n$ with smaller γ is observed to excel in detecting a change of scale parameters, while the power regarding the location parameter change is shown to increase when incrementing γ. We speculate that this phenomenon occurs as the increase in volatility can contribute to triggering the normalizing factor, $\hat{\tau}_n$ of (8), to become excessively large when employing larger γ, and thereby lower the values of the test statistics $\hat{T}_n$. Hence, if one intends to utilize $\hat{T}_n$ in practice, one can first fit a suitable location-scale model like the ARMA-GARCH model to a given time series, then choose a suitable γ accordingly to the obtained results. In case one desires to detect a change in the location parameters that concurrently appear to be statistically significant, one can employ a large γ, possibly around 1, to balance between the detection ability of both location and scale parameters. If not, a small γ would be adequate to maximize the detection ability of quantile changes induced from a scale parameter change.

Thus far, our findings have demonstrated the respective strengths of $\hat{T}_n$, $\tilde{T}_n$ and $\check{T}_n$. For both pure GARCH-type and ARMA-GARCH models, $\hat{T}_n$ retains its superiority in terms of the size, therefore is more robust against false alarms. Moreover, for GARCH-type models, $\hat{T}_n$ with smaller γ, namely $\gamma = 0.2$, $\tilde{T}_{n,2}$, and $\check{T}_{n,2}$ outperform the remaining tests in terms of power, while for ARMA-type models, $\tilde{T}_n$ and $\check{T}_n$ tend to outperform the others. In conclusion, we recommend using both $\hat{T}_n$ with small γ and $\check{T}_n$ in practice. This is because $\hat{T}_n$ with small γ complements the drawback of both $\tilde{T}_n$ and $\check{T}_n$, also performing well in handling scale parameter changes, while $\check{T}_n$ has strong detection ability in general, being more stable than $\tilde{T}_n$. Those two tests have their own merits that are more adaptable to a specific situation, which is further reflected in our real data analysis addressed below.

5 Data analysis

In this section, we analyze the daily returns of the Brent Oil futures from September 1, 2010 to September 28, 2018 (2088 observations). The raw indices are depicted in Fig. 1a. These datasets are available at the website "investing.com," named "Oil"

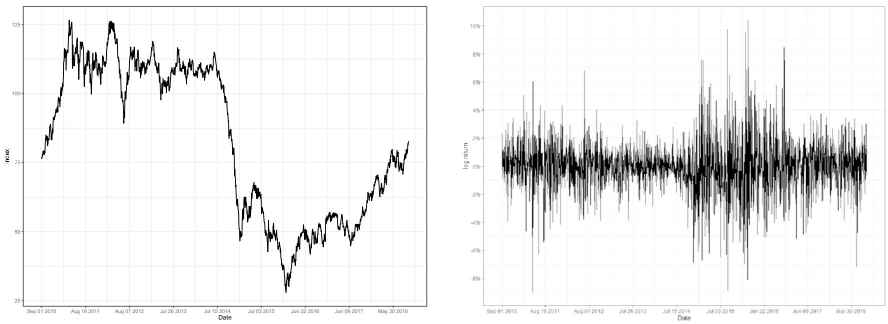
Table 1 p values for the backtesting results of the conditional quantile estimates regarding the tail quantiles $\tau \in \{0.95, 0.9, 0.1, 0.05\}$ of Oil

	$\tau = 0.95$	$\tau = 0.9$	$\tau = 0.1$	$\tau = 0.05$
Prior				
Kupiec	0.961	0.943	0.943	0.818
Christoffersen	0.025	0.067	0.818	0.919
Posterior				
Kupiec	0.809	0.809	0.930	0.952
Christoffersen	0.094	0.673	0.099	0.082

Table 2 Results of the change point detection test for tail quantiles $\tau \in \{0.95, 0.9, 0.1, 0.05\}$ regarding Oil

	$\hat{T}_n$		$\check{T}_n$
	$\gamma = 0.2$	$\gamma = 1$	
$\tau = 0.95$	Jan/15/15	Jan/15/15	Sep/29/14
$\tau = 0.9$	Jan/13/15	–	Sep/29/14
$\tau = 0.1$	Aug/11/14	–	Oct/10/14
$\tau = 0.05$	Sep/26/14	–	Oct/10/14

"–" indicates that the test did not detect any significant change point

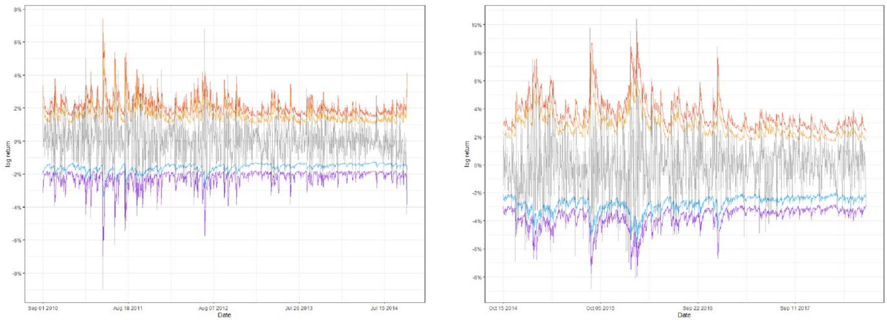


(a) Raw price index (b) log-returns

Fig. 1 Plot of (**a**) the price index of Oil and (**b**) its log-returns. A noticeable drop in price is observable during the late 2014

for simplicity, and the log returns $r_t = 100 \times \{\log(y_t) - \log(y_{t-1})\}$ $(t \geq 2)$ are used in our analysis.

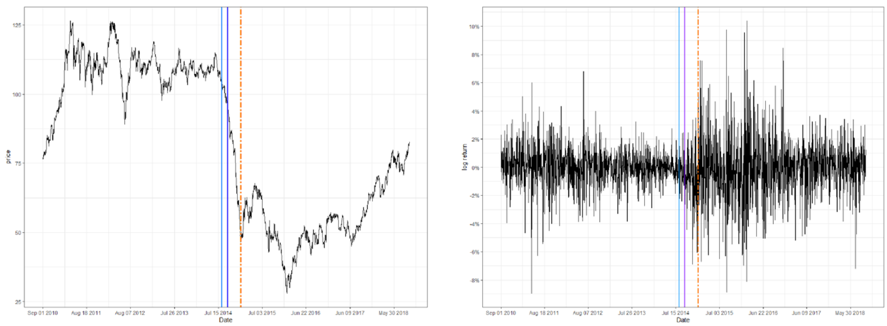
We first fit an AR(1)-GARCH(1, 1) model to Oil, then notice that the AR parameter is revealed to be significant according to a test. We then obtain the conditional quantile regression estimators presented in (5) for $\tau \in \{0.95, 0.9, 0.1, 0.05\}$, which are depicted in Fig. 1b, and examine whether a change in tail quantiles is present for those τ's. For this purpose, we employ $\hat{T}_n$ with $\gamma = 0.2$ and 1, alongside with $\check{T}_n$, to cope with a change regarding both the AR parameter and the conditional volatility, see Sect. 4.2 for the discussion of the respective strengths of the test statistics.



(a) Before (b) After

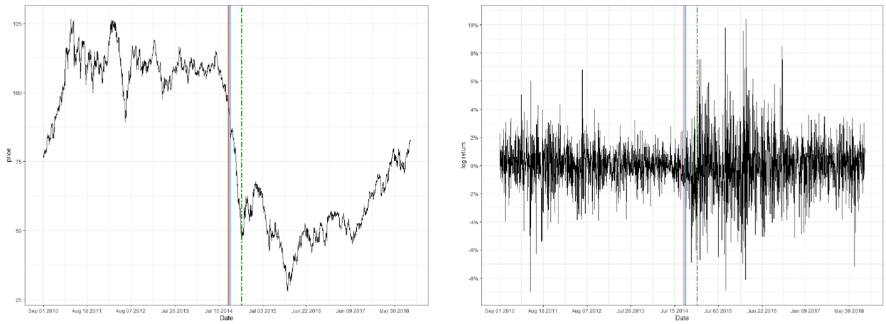
Fig. 2 Plot of the estimated conditional quantiles of Oil (**a**) before and (**b**) after the change point at $\tau = 0.95$ (red), 0.9 (orange), 0.1 (blue), and 0.05 (purple) (color figure online)

Table 2 and Figs. 3 and 4 illustrate the results for the change point test applied to Oil. They show that $\check{T}_n$ detects a change on September 29, 2014 and October 10, 2014 for the upper and lower quantiles, respectively. Similarly, $\hat{T}_n$ with $\gamma = 0.2$ is shown to detect a change on January 15, 2015 and two days prior, respectively, for upper quantiles $\tau = 0.95$ and 0.9, and also on August 11, 2014 and September 26, 2014 for lower quantiles $\tau = 0.1$ and 0.05, respectively. Moreover, it is shown that $\hat{T}_n$ with $\gamma = 1$ detects a change in the 95% quantile on January 15, 2015, but detects no significant change in other tail quantiles. These results signify that no quantile changes are induced from the AR parameter change, except for the uppermost quantile, and with a high possibility, a change may exist in the conditional quantile originated from the conditional volatility change. Indeed, Fig. 3 displays that an increase in volatility exists, including that of the tail quantiles, around the locations of the detected change point.



(a) raw price index (b) log-return

Fig. 3 Location of the change points of Oil at $\tau \in \{0.95, 0.9, 0.1, 0.05\}$, using $\hat{T}_n$ with $\gamma = 0.2$. Red ($\tau = 0.95$) and orange ($\tau = 0.9$) dotted lines symbolize the detected location of upper quantiles, while blue ($\tau = 0.1$) and purple ($\tau = 0.05$) solid lines denote those of the lower quantiles (color figure online)



(a) raw price index (b) log-return

Fig. 4 Location of the change points of Oil at $\tau \in \{0.95, 0.9, 0.1, 0.05\}$, using $\hat{T}_n$ with $\gamma = 1$ and $\check{T}_n$. Green dotted lines symbolize the detected location of the former test, while red ($\tau = 0.95, 0.9$) and blue ($\tau = 0.1, 0.05$) solid lines denote those of the latter test (color figure online)

Moreover, to further analyze the model structure prior and posterior the change point, we first divide Oil into two separate subseries, based on the results of $\check{T}_n$, then independently apply the two backtesting measures introduced in Sect. 4.1. Table 1 and Fig. 2, respectively, depict the results of the two measures and the estimated conditional quantiles of Oil for the two subseries. Although the result mostly indicates the adequacy of the quantile regression estimates across all tail quantiles, Christoffersen's test with $\tau = 0.95$ applied to Oil before the change point appears to reject the null at the significance level of 0.05. This is actually because the first subseries of Oil prior to the change point also contains a change point of a smaller magnitude on November 7, 2012 according to $\check{T}_n$: a decrease in volatility can be observed in Fig. 3b.

The results suggest that $\hat{T}_n$ can be preferable to $\check{T}_n$ when the underlying structure of the dataset is heavily asymmetric. In fact, Fig. 1a displays a steep plunge in the price of the futures from September 29, 2014 to January 13, 2015. This is reasoned to cause the lower quantiles of the log returns to depart from their original structure prior to the upper quantiles, which is concurrent with the results of $\hat{T}_n$ mentioned earlier. Moreover, $\hat{T}_n$ with $\gamma = 0.2$ detecting a change a month prior to this event implicates that there could have been a warning for this structural break during July and August, 2014, as the actual price drop commenced around July 15, 2014, as seen in Fig. 1a.

Our findings reaffirm the merits of $\hat{T}_n$ as it can detect conditional quantile changes owing to location and scale parameter changes by adjusting γ properly, particularly those induced from volatility changes. Moreover, $\check{T}_n$ can simultaneously detect changes in all aspects and is advantageous when one aims to detect a small change in tail volatilities, see Sect. 4.2. We in general recommend using $\hat{T}_n$ with small γ when the location parameters of the given time series is insignificant, and $\check{T}_n$ otherwise, as each test has its own merit to react more sensitively to a given situation. Overall, our empirical study strongly confirms the validity of our tests.

6 Concluding remarks

In this paper, we have proposed the three types of CUSUM tests to detect a change in conditional quantiles for a general location-scale time series models, derived their limiting null distributions, and explored the consistency under the alternative. Our simulation study has demonstrated that the proposed CUSUM tests in general performs stable and powerful under a variety of model settings, and also have exhibited that $\hat{T}_n$, $\tilde{T}_n$ and $\check{T}_n$ have complementary relationships, namely, $\hat{T}_n$ is more geared toward being more stable, while $\tilde{T}_n$ and $\check{T}_n$ are more suitable when detecting a change of a small magnitude. A real data analysis has also been conducted for illustration using the log returns of Brent oil futures, which not only demonstrated the validity of our proposed tests but also additionally revealed the adaptability and versatility of $\hat{T}_n$. All these results strongly confirm the functionality of the proposed tests.

Acknowledgements We sincerely thank the Editor, an AE, and anonymous reviewers for their careful reading and valuable comments that improve the quality of the paper. This research is supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (No. 2021R1A2C1004009).

Appendix

In this section, we provide the proofs of some results in Sects. 2 and 3.

Proof of Theorem 1 It suffices to prove (9). Put

$$\Lambda_t(\vartheta) = (\tilde{g}_t(\mu) - g_t(\mu_0))/h_t^{1/2}(\theta_0) + \xi(\tilde{h}_t(\theta)/h_t(\theta_0))^{1/2} - \xi_0.$$

Noting that $\psi_\tau(y_t - \hat{q}_t) = \tau - I(\eta_t \le \xi_0 + \Lambda_t(\hat{\vartheta}_n))$, we can express

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \max_{1\le k\le n} \Bigg| \sum_{t=1}^{k} |y_{t-1}|^\gamma \Big\{ \psi_\tau(y_t - \hat{q}_t) - \psi_\tau(y_t - q_t(\vartheta_0)) \Big\} \\
&\qquad - \frac{k}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma \Big\{ \psi_\tau(y_t - \hat{q}_t) - \psi_\tau(y_t - q_t(\vartheta_0)) \Big\} \Bigg| \\
&\le \frac{2}{\sqrt{n}} \max_{1\le k\le n} \Bigg| \sum_{t=1}^{k} |y_{t-1}|^\gamma \Big\{ I(\eta_t \le \xi_0 + \Lambda_t(\hat{\vartheta}_n)) - F_\eta(\xi_0 + \Lambda_t(\hat{\vartheta}_n)) + F_\eta(\xi_0) - I(\eta_t \le \xi_0) \Big\} \Bigg| \\
&\qquad + \frac{1}{\sqrt{n}} \max_{1\le k\le n} \Bigg| \sum_{t=1}^{k} |y_{t-1}|^\gamma \Big\{ F_\eta(\xi_0 + \Lambda_t(\hat{\vartheta}_n)) - F_\eta(\xi_0) \Big\} \\
&\qquad - \frac{k}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma \Big\{ F_\eta(\xi_0 + \Lambda_t(\hat{\vartheta}_n)) - F_\eta(\xi_0) \Big\} \Bigg| \\
&:= 2A_n + B_n.
\end{aligned}$$

Owing to Proposition 1, for any $\delta \in (0,1)$, there exists a (large enough) $L > 0$ such that $P(\hat{\vartheta}_n \in \mathcal{N}_{L/\sqrt{n}}) \ge 1 - \delta$, where $\mathcal{N}_{L/\sqrt{n}}$ is a compact neighborhood of ϑ_0 with

$||\vartheta - \vartheta_0|| \le L/\sqrt{n}$ for all $\vartheta \in \mathcal{N}_{L/\sqrt{n}}$. Given any fixed $\zeta > 0$, we decompose $\mathcal{N}_{L/\sqrt{n}}$ into a finite number of subsets $\mathcal{D}_1, \ldots, \mathcal{D}_N$ for some $N = N(\zeta) \ge 1$, with their diameters less than $\zeta/\sqrt{n}$. We then choose points ϑ_j from $\mathcal{D}_j$. Then, in case of $\hat{\vartheta}_n \in \mathcal{D}_j$, we have

$$I(\eta_t \le \xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^-) \le I(\eta_t \le \xi_0 + \Lambda_t(\hat{\vartheta}_n)) \le I(\eta_t \le \xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^+)$$

with $\Lambda_{tj}^- = \inf_{\vartheta\in\mathcal{D}_j} \Lambda_t(\vartheta) - \Lambda_t(\vartheta_j)$ and $\Lambda_{tj}^+ = \sup_{\vartheta\in\mathcal{D}_j} \Lambda_t(\vartheta) - \Lambda_t(\vartheta_j)$. Putting

$$\begin{aligned} A_{kj}^+ =& \frac{1}{\sqrt{n}} \sum_{t=1}^{k} |y_{t-1}|^\gamma \Big\{ I(\eta_t \le \xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^+) \\ &- F_\eta(\xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^+) - F_\eta(\xi_0) + I(\eta_t \le \xi_0) \Big\}, \\ A_{kj}^- =& \frac{1}{\sqrt{n}} \sum_{t=1}^{k} |y_{t-1}|^\gamma \Big\{ I(\eta_t \le \xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^-) \\ &- F_\eta(\xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^-) - F_\eta(\xi_0) + I(\eta_t \le \xi_0) \Big\}, \end{aligned}$$

and using (A5), (A6), and the mean value theorem, we can show that

$$A_n \le \max_{1\le j\le N} \max_{1\le k\le n} |A_{kj}^+| + \max_{1\le j\le N} \max_{1\le k\le n} |A_{kj}^-| + r_n \tag{12}$$

with

$$\begin{aligned} r_n =& \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \max_{1\le j\le N} \sup_{\vartheta\in\mathcal{D}_j} |y_{t-1}|^\gamma |F_\eta(\xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^+) - F_\eta(\xi_0 + \Lambda_t(\vartheta))| \\ &+ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \max_{1\le j\le N} \sup_{\vartheta\in\mathcal{D}_j} |y_{t-1}|^\gamma F_\eta(\xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^-) - F_\eta(\xi_0 + \Lambda_t(\vartheta))| \\ =& \zeta O_P(1), \end{aligned} \tag{13}$$

which can be made arbitrarily small by taking a small ζ.

Next, putting

$$\begin{aligned} e_{tj}^+ =& |y_{t-1}|^\gamma \Big\{ I(\eta_t \le \xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^+) \\ &- F_\eta(\xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^+) - F_\eta(\xi_0) + I(\eta_t \le \xi_0) \Big\}, \\ e_{tj}^- =& |y_{t-1}|^\gamma \Big\{ I(\eta_t \le \xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^-) \\ &- F_\eta(\xi_0 + \Lambda_t(\vartheta_j) + \Lambda_{tj}^-) - F_\eta(\xi_0) + I(\eta_t \le \xi_0) \Big\}, \end{aligned}$$

we set $S_{kj}^+ = \sum_{t=1}^k e_{tj}^+$ and $S_{kj}^- = \sum_{t=1}^k e_{tj}^-$, Moreover, for $B > 0$, we set

$$\tilde{S}^{+}_{kj} = \sum_{t=1}^{k} \tilde{e}^{+}_{tj}, \quad \tilde{S}^{-}_{kj} = \sum_{t=1}^{k} \tilde{e}^{-}_{tj}$$

with

$$\begin{aligned} \tilde{e}^{+}_{tj} =& e^{+}_{tj} I\left(\sum_{s=1}^{t} |y_s|^{2\gamma} \sup_{\vartheta\in\Theta} \left\| \frac{\partial \Lambda_s(\vartheta)}{\partial\vartheta} \right\| \le Bn \right), \\ \tilde{e}^{-}_{tj} =& e^{-}_{tj} I\left(\sum_{s=1}^{t} |y_s|^{2\gamma} \sup_{\vartheta\in\Theta} \left\| \frac{\partial \Lambda_s(\vartheta)}{\partial\vartheta} \right\| \le Bn \right). \end{aligned}$$

Then, we get martingale arrays $\{(\tilde{S}^{+}_{kj}, \mathcal{F}_t); k = 1, \ldots, n\}$ and $\{(\tilde{S}^{-}_{kj}, \mathcal{F}_t); k = 1, \ldots, n\}$, where $\mathcal{F}_t = \sigma(\eta_t, \eta_{t-1}, \ldots)$.

Note that $P(\tilde{e}^{+}_{tj} \neq e^{+}_{tj}$ for some $t = 1, \ldots, n)$ can be made arbitrarily small by taking a large B, which can be seen using the mean value theorem, Hölder's inequality, (A5) and (A6). Then, for any $\zeta', \lambda > 0$, $0 < \delta < 1$ and $p > 1$, using the sub-martingale inequality and taking a sufficiently large $B := B(\zeta)$, we can get

$$\begin{aligned} &\limsup_{n\to\infty} P(A^{+}_{n} \ge \lambda) \\ &\le \limsup_{n\to\infty} P\left(\max_{1\le j\le N} \max_{1\le k\le n} |A^{+}_{kj}| \ge \lambda \right) \le \limsup_{n\to\infty} \sum_{j=1}^{N} P\left(\max_{n^{\delta}\le k\le n} |S^{+}_{kj}| \ge \sqrt{n}\lambda \right) \\ &\le \limsup_{n\to\infty} \sum_{j=1}^{N} P\left(\max_{1\le k\le n} |\tilde{S}^{+}_{kj}| \ge \sqrt{n}\lambda \right) + \zeta' \le \limsup_{n\to\infty} \sum_{j=1}^{N} E\tilde{S}^{+2p}_{nj} / n^{p} \lambda^{2p} + \zeta'. \end{aligned} \tag{14}$$

Moreover, using Rosenthal's inequality (Hall and Heyde 1980), we get

$$E\tilde{S}^{+2p}_{nj} \le C\left(E\left[\sum_{t=1}^{n} E(\tilde{e}^{+2}_{tj} | \mathcal{F}_{t-1}) \right]^{p} + \sum_{t=1}^{n} E(\tilde{e}^{+2p}_{tn}) \right), \ C > 0,$$

and further,

$$E\left[\sum_{t=1}^{n} E(\tilde{e}^{+2}_{tj} | \mathcal{F}_{t-1}) \right]^{p} = O(n^{p/2}) \quad \text{and} \quad \sum_{t=1}^{n} E(\tilde{e}^{+2p}_{tj}) = O(n),$$

which leads to $\sum_{j=1}^{N} E\tilde{S}^{+2p}_{nj} = O(n^{p/2} + n)$. Then, in view of (14), we obtain

$$\max_{1\le j\le N} \max_{1\le k\le n} |A^{+}_{kj}| = o_P(1).$$

Analogously, we have $\max_{1\le j\le N} \max_{1\le k\le n} |A^{-}_{kj}| = o_P(1)$. Hence, from (12) and (13), we have $A_n = o_P(1)$.

Next, we handle B_n. Set

$$B_{nk} = \left| \sum_{t=1}^{k} |y_{t-1}|^\gamma \left\{ F_\eta(\xi_0 + \Lambda_t(\hat{\vartheta}_n)) - F_\eta(\xi_0) \right\} \right.$$
$$\left. - \frac{k}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma \left\{ F_\eta(\xi_0 + \Lambda_t(\hat{\vartheta}_n)) - F_\eta(\xi_0) \right\} \right|.$$

Note that for $k_n = n^{1/3}$, $\frac{1}{\sqrt{n}} \max_{1 \le k \le k_n} B_{nk} = o_P(1)$. Thus it suffices to show that

$$B_n' := \frac{1}{\sqrt{n}} \max_{k_n \le k \le n} B_{nk} = o_P(1).$$

To this end, we use Taylor's theorem to get

$$F_\eta(\xi_0 + \Lambda_t(\hat{\vartheta}_n)) - F_\eta(\xi_0) = \Lambda_t(\hat{\vartheta}_n) f_\eta(\xi_0) + \frac{1}{2} \Lambda_t^2(\hat{\vartheta}_n) f_\eta'(\xi_t^*)$$

for some number ξ_t^* lying between ξ_0 and $\xi_0 + \Lambda_t(\hat{\vartheta}_n)$. Then, using this, (A2), (A5), (A6), and our assumption on f_η', we can show that for some $K > 0$,

$$\begin{aligned}
B_n' \le & K\sqrt{n}||\hat{\mu}_n - \mu_0|| \max_{k_n \le k \le n} \left\| \frac{1}{k} \sum_{t=1}^{k} |y_{t-1}|^\gamma \frac{\partial g_t(\mu_0)}{\partial \mu} \frac{1}{h_t^{1/2}(\theta_0)} \right. \\
& \left. - \frac{1}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma \frac{\partial g_t(\mu_0)}{\partial \mu} \frac{1}{h_t^{1/2}(\theta_0)} \right\| \\
& + K\sqrt{n}||\hat{\theta}_n - \theta_0|| \max_{k_n \le k \le n} \left\| \frac{1}{k} \sum_{t=1}^{k} |y_{t-1}|^\gamma \frac{\partial h_t(\theta_0)}{\partial \theta} \frac{1}{\sqrt{2} h_t(\theta_0)} \right. \\
& \left. - \frac{1}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma \frac{\partial h_t(\theta_0)}{\partial \theta} \frac{1}{\sqrt{2} h_t(\theta_0)} \right\| \\
& + K\sqrt{n}|\hat{\xi}_n - \xi_0| \max_{k_n \le k \le n} \left| \frac{1}{k} \sum_{t=1}^{k} |y_{t-1}|^\gamma - \frac{1}{n} \sum_{t=1}^{n} |y_{t-1}|^\gamma \right| + o_P(1) \\
= & o_P(1),
\end{aligned}$$

due to the ergodicity of $\{y_t\}$, which entails $B_n = o_P(1)$. As we can check $\hat{\tau}_1^2 = \tau_1^2 + o_P(1)$, we finally get $\hat{T}_n - T_n = o_P(1)$, which in turn implicates the weak convergence of $\hat{T}_n$ to $\sup_{0 \le s \le 1} |B^\circ(s)|$. □

Proof of Theorem 4 For $i = 1, 2$, we set

$$\Lambda_{ti}^*(\vartheta) = \frac{\tilde{g}_{ti}(\mu) - g_{ti}(\mu_0^*)}{h_{ti}^{1/2}(\theta_i)} + \frac{\xi \tilde{h}_{ti}^{1/2}(\theta) - \xi_0^* h_{ti}^{1/2}(\theta_0^*)}{h_{ti}^{1/2}(\theta_i)},$$
$$\xi_{ti} = \frac{g_{ti}(\mu_0^*) - g_{ti}(\mu_i) + \xi_0^* h_{ti}^{1/2}(\theta_0^*)}{h_{ti}^{1/2}(\theta_i)}.$$

Noting that $\psi_\tau(y_{t1} - \hat{q}_{t1}) = \tau - I(\eta_t \le \Lambda_t^*(\hat{\vartheta}_n) + \xi_{t1})$, we can express for $k \le k_0$,

$$\frac{1}{n}\sum_{t=1}^{k} |y_{t-1,1}|^\gamma \psi_\tau(y_{t1} - \hat{q}_{t1}) = \frac{1}{n}\sum_{t=1}^{k} |y_{t-1,1}|^\gamma \psi_\tau(y_{t1} - q_{t1}(\vartheta_0^*)) + I_{nk} + II_{nk},$$

where

$$I_{nk} = -\frac{1}{n}\sum_{t=1}^{k} |y_{t-1,1}|^\gamma \Big\{ I(\eta_t \le \Lambda_{t1}^*(\hat{\vartheta}_n) + \xi_{t1}) - F_\eta(\Lambda_{t1}^*(\hat{\vartheta}_n) + \xi_{t1}) + F_\eta(\xi_{t1}) - I(\eta_t \le \xi_{t1}) \Big\},$$
$$II_{nk} = -\frac{1}{n}\sum_{t=1}^{k} |y_{t-1,1}|^\gamma \Big\{ F_\eta(\Lambda_{t1}^*(\hat{\vartheta}_n) + \xi_{t1}) - F_\eta(\xi_{t1}) \Big\}.$$

Then, following the lines similar to those used for verifying $A_n = o_P(1)$ in the proof of Theorem 1, with $\mathcal{N}_{L/\sqrt{n}}$, $\Lambda_t(\vartheta)$, and ξ_0, therein, replaced by $\mathcal{N}_L^* = \{\vartheta : ||\vartheta - \vartheta_0^*|| \le L\}$, $\Lambda_{t1}^*(\vartheta)$, and ξ_{t1}, respectively, where L can be taken to be arbitrarily small, $\mathcal{N}_L^*$ and L play the role of the partition $\mathcal{D}_1$ and ζ (namely, $\mathcal{N}_L^* = \mathcal{D}_1$, $L = \zeta$, and $N(\zeta) = 1$), and both A_{k1}^+ and A_{k1}^- are accordingly reformulated with the denominator $\sqrt{n}$ replaced by n, we can similarly verify that $\max_{1\le k\le k_0} |I_{nk}| = o_P(1)$. Also, using (A2), (A5), (A6) (modified with $g_{ti}, h_{ti}, \tilde{g}_{ti}, \tilde{h}_{ti}$), (B1), the ergodicity, and the mean value theorem, we can readily check that $\max_{1\le k\le k_0} |II_{nk}| = o_P(1)$. Subsequently, we get

$$\frac{1}{n}\sum_{t=1}^{k} |y_{t-1,1}|^\gamma \psi_\tau(y_{t1} - \hat{q}_{t1}) = \frac{1}{n}\sum_{t=1}^{k} |y_{t-1,1}|^\gamma \psi_\tau(y_{t1} - q_{t1}(\vartheta_0^*)) + o_P(1), \tag{15}$$

uniformly in $k \le k_0$. Similarly, additionally harnessing (11), we can verify

$$\frac{1}{n}\sum_{t=k_0+1}^{k}|y_{t-1,2}|^{\gamma}\psi_{\tau}(y_{t2}-\hat{q}_{t2}) = \frac{1}{n}\sum_{t=k_0+1}^{n}|y_{t-1,2}|^{\gamma}\psi_{\tau}(y_{t2}-q_{t2}(\vartheta_0^*)) + o_P(1), \tag{16}$$

uniformly in $k > k_0$. Therefore, we can have

$$\begin{aligned}
\frac{|\hat{\tau}_n\hat{T}_n|}{\sqrt{n}} \geq \frac{1}{n}\Big|&\sum_{t=1}^{k_0}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{t1}-\hat{q}_{t1}) \\
&- \frac{k_0}{n}\Big\{\sum_{t=1}^{k_0}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{t1}-\hat{q}_{t1}) + \sum_{t=k_0+1}^{n}|y_{t-1,2}|^{\gamma}\psi_{\tau}(y_{t2}-\hat{q}_{t2})\Big\}\Big| \\
= \frac{1}{n}\Big|&\sum_{t=1}^{k_0}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{t1}-q_{t1}(\vartheta_0^*)) \\
&- \frac{k_0}{n}\Big\{\sum_{t=1}^{k_0}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{ti}-q_{t1}(\vartheta_0^*)) \\
&+ \sum_{t=k_0+1}^{n}|y_{t-1,2}|^{\gamma}\psi_{\tau}(y_{t2}-q_{t2}(\vartheta_0^*))\Big\}\Big| + o_P(1) \\
&\xrightarrow{P} \kappa_0(1-\kappa_0)|d_1-d_2| > 0,
\end{aligned}$$

due to the ergodicity and (B2), so that $\hat{T}_n \to \infty$ in probability as $\hat{\tau}_n = \tau_0 + o_P(1)$ for some $\tau_0 > 0$.

Moreover, putting

$$\begin{aligned}
\hat{\mathcal{L}}_n(k) = \frac{1}{n}\Big|&\sum_{t=1}^{k}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{t1}-\hat{q}_{t1}) \\
&- \frac{k}{n}\Big\{\sum_{t=1}^{k_0}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{t1}-\hat{q}_{t1}) + \sum_{t=k_0+1}^{n}|y_{t-1,2}|^{\gamma}\psi_{\tau}(y_{t2}-\hat{q}_{t2})\Big\}\Big|,
\end{aligned}$$

we can see that $\max_{0\leq s\leq 1}|\hat{\mathcal{L}}_n([ns]) - \mathcal{L}_n([ns])| = o_P(1)$ by virtue of (15) and (16), where

$$\begin{aligned}
\mathcal{L}_n(k) = \frac{1}{n}\Big|&\sum_{t=1}^{k}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{t1}-q_{t1}(\vartheta_0^*))I(k \leq k_0) \\
&+ \Big\{\sum_{t=1}^{k_0}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{t1}-q_{t1}(\vartheta_0^*)) + \sum_{t=k_0+1}^{k}|y_{t-1,2}|^{\gamma}\psi_{\tau}(y_{t2}-q_{t2}(\vartheta_0^*))\Big\}I(k > k_0) \\
&- \frac{k}{n}\Big\{\sum_{t=1}^{k_0}|y_{t-1,1}|^{\gamma}\psi_{\tau}(y_{ti}-q_{t1}(\vartheta_0^*)) + \sum_{t=k_0+1}^{n}|y_{t-1,2}|^{\gamma}\psi_{\tau}(y_{ti}-q_{t2}(\vartheta_0^*))\Big\}\Big|.
\end{aligned}$$

Furthermore, using the ergodicity, we can easily check that $\sup_{0\le s\le 1} |\mathcal{L}_n([ns]) - \mathcal{L}(s)| = o_P(1)$, where

$$\mathcal{L}(s) = s(1-\kappa_0)|d_1 - d_2| I(s \le \kappa_0) + \kappa_0(1-s)|d_1 - d_2| I(\kappa_0 < s \le 1),$$

which has a maximum value $\kappa_0(1-\kappa_0)|d_1 - d_2|$ uniquely at $s = \kappa_0$. This with the argmax theorem yields $\hat{k}_n/n \to \kappa_0$ in probability, which validates the theorem. □

On the issue regarding (B1) in Remark 4. For $t \le k_0$, put

$$\Lambda_t^\dagger(\vartheta) = (\tilde{g}_{t1}(\mu) - g_{t1}(\mu_1))/h_{t1}^{1/2}(\theta_1) + \xi(\tilde{h}_{t1}(\theta)/h_{t1}(\theta_1))^{1/2},$$
$$\Lambda_t^\circ(\vartheta) = (g_{t1}(\mu) - g_{t1}(\mu_1))/h_{t1}^{1/2}(\theta_1) + \xi(h_{t1}(\theta)/h_{t1}(\theta_1))^{1/2}.$$

Then, we can express

$$\rho_\tau(y_t - \tilde{q}_t(\vartheta)) = (y_{t1} - \tilde{q}_{t1}(\vartheta))\psi_\tau(y_{t1} - \tilde{q}_{t1}(\vartheta)),$$
$$\rho_\tau(y_t - q_t(\vartheta)) = (y_{t1} - q_{t1}(\vartheta))\psi_\tau(y_{t1} - q_{t1}(\vartheta))$$

with

$$\psi_\tau(y_t - \tilde{q}_t(\vartheta)) = \tau - I(\eta_t \le \Lambda_t^\dagger(\vartheta)), \quad \psi_\tau(y_t - q_{t1}(\vartheta)) = \tau - I(\eta_t \le \Lambda_t^\circ(\vartheta)).$$

We first verify

$$\sup_{\vartheta\in\Theta^*} \left| \frac{1}{n}\sum_{t=1}^{k_0} \rho_\tau(y_t - \tilde{q}_t(\vartheta)) - \frac{1}{n}\sum_{t=1}^{k_0} \rho_\tau(y_{t1} - q_{t1}(\vartheta)) \right| = o_P(1). \tag{17}$$

Due to (A5) and (A6), it suffices to show that

$$\sup_{\vartheta\in\Theta^*} \left| \frac{1}{n}\sum_{t=1}^{k_0} (y_t - \tilde{q}_t(\vartheta)) I(\eta_t \le \Lambda_t^\dagger(\vartheta)) - \frac{1}{n}\sum_{t=1}^{k_0} (y_{t1} - q_{t1}(\vartheta)) I(\eta_t \le \Lambda_t^\circ(\vartheta)) \right| = o_P(1),$$

or equivalently, for $\delta \in (0,1)$,

$$\sup_{\vartheta\in\Theta^*} \left| \frac{1}{n}\sum_{t=n^\delta}^{k_0} (y_t - q_t(\vartheta)) I(\eta_t \le \Lambda_t^\dagger(\vartheta)) - \frac{1}{n}\sum_{t=n^\delta}^{k_0} (y_{t1} - q_{t1}(\vartheta)) I(\eta_t \le \Lambda_t^\circ(\vartheta)) \right| = o_P(1),$$

which we can check to hold via using the Cauchy–Schwarz inequality, provided

$$\sup_{\vartheta\in\Theta^*} \frac{1}{n}\sum_{t=n^\delta}^{k_0} (I(\eta_t \le \Lambda_t^\dagger(\vartheta)) - I(\eta_t \le \Lambda_t^\circ(\vartheta))^2 = o_P(1). \tag{18}$$

To verify (18), we partition Θ^* into $\mathcal{D}_j$ with diameter less than ζ, $j = 1, \ldots, N = N(\zeta)$, and choose points ϑ_j from $\mathcal{D}_j$. Then, in case of $\vartheta \in \mathcal{D}_j$, we have

$$I(\eta_t \le \Lambda_t^\circ(\vartheta_j) + \Lambda_{tj}^{\dagger(1)}) \le I(\eta_t \le \Lambda_t^\dagger(\vartheta)) \le I(\eta_t \le \Lambda_t^\circ(\vartheta_j) + \Lambda_{tj}^{\dagger(2)})$$

with $\Lambda_{tj}^{\dagger(1)} = \inf_{\vartheta\in\mathcal{D}_j} \Lambda_t^\dagger(\vartheta) - \Lambda_t^\circ(\vartheta_j)$ and $\Lambda_{tj}^{\dagger(2)} = \sup_{\vartheta\in\mathcal{D}_j} \Lambda_t^\dagger(\vartheta) - \Lambda_t^\circ(\vartheta_j)$. Also,

$$I(\eta_t \le \Lambda_t^\circ(\vartheta_j) + \Lambda_{tj}^{\circ(1)}) \le I(\eta_t \le \Lambda_t^\circ(\vartheta)) \le I(\eta_t \le \Lambda_t^\circ(\vartheta_j) + \Lambda_{tj}^{\circ(2)})$$

with $\Lambda_{tj}^{\circ(1)} = \inf_{\vartheta\in\mathcal{D}_j} \Lambda_t^\circ(\vartheta) - \Lambda_t^\circ(\vartheta_j)$ and $\Lambda_{tj}^{\circ(2)} = \sup_{\vartheta\in\mathcal{D}_j} \Lambda_t^\circ(\vartheta) - \Lambda_t^\circ(\vartheta_j)$.

By simple algebras using (A5) and (A6), it can be seen that

$$\begin{aligned} C_n :=& \max_{1\le j\le N} \max_{1\le k\ne l\le 2} \frac{1}{n} \sum_{t=n^\delta}^{k_0} \left[I(\eta_t \le \Lambda_t^\circ(\vartheta_j) + \Lambda_{tj}^{\dagger(k)}) - I(\eta_t \le \Lambda_t^\circ(\vartheta_j) + \Lambda_{tj}^{\circ(l)}) \right]^2 \\ =& \zeta O_P(1) + o_P(1). \end{aligned}$$

As ζ can be taken to be arbitrarily small, C_n becomes $o_P(1)$, which ensures (18) and thereby (17) as the quantity in (18) is no more than C_n. Moreover, following the lines similar to the above, we can also check

$$\sup_{\vartheta\in\Theta^*} \left| \frac{1}{n} \sum_{t=1}^{k_0} \rho_\tau(y_{t1} - q_{t1}(\vartheta)) - \kappa_0 E\rho_\tau(y_{t1} - q_{t1}(\vartheta)) \right| = o_P(1),$$

and thus, due to (17),

$$\sup_{\vartheta\in\Theta^*} \left| \frac{1}{n} \sum_{t=1}^{k_0} \rho_\tau(y_t - \tilde{q}_t(\vartheta)) - \kappa_0 E\rho_\tau(y_{t1} - q_{t1}(\vartheta)) \right| = o_P(1). \tag{19}$$

Similarly, it can be shown that

$$\begin{aligned} &\sup_{\vartheta\in\Theta^*} \left| \frac{1}{n} \sum_{t=k_0+1}^{n} \rho_\tau(y_t - \tilde{q}_t(\vartheta)) - (1-\kappa_0)\rho_\tau(y_{t2} - q_{t2}(\vartheta)) \right| \\ &\le \sup_{\vartheta\in\Theta^*} \left| \frac{1}{n} \sum_{t=k_0+1}^{n} \rho_\tau(y_t - \tilde{q}_t(\vartheta)) - \frac{1}{n} \sum_{t=k_0+1}^{n} \rho_\tau(y_{t2} - q_{t2}(\vartheta)) \right| \\ &\quad + \sup_{\vartheta\in\Theta^*} \left| \frac{1}{n} \sum_{t=k_0+1}^{n} \rho_\tau(y_{t2} - q_{t2}(\vartheta)) - (1-\kappa_0) E\rho_\tau(y_{t2} - q_{t2}(\vartheta)) \right| \\ &= o_P(1), \end{aligned} \tag{20}$$

then combining (19) and (20), we have

$$\begin{aligned} &\sup_{\vartheta\in\Theta^*} \left| \frac{1}{n} \sum_{t=1}^{n} \rho_\tau(y_t - \tilde{q}_t(\vartheta)) - \left\{ \kappa_0 E\rho_\tau(y_{t1} \right.\right. \\ &\quad \left.\left. - q_{t1}(\vartheta)) + (1-\kappa_0) E\rho_\tau(y_{t2} - q_{t2}(\vartheta)) \right\} \right| = o_P(1), \end{aligned}$$

which validates the conclusion in Remark 4. Namely, (B1) holds true if $\kappa_0 E\rho_\tau(y_{t1} - q_{t1}(\vartheta)) + (1 - \kappa_0)E\rho_\tau(y_{t2} - q_{t2}(\vartheta))$ has its maximum value uniquely at an interior point ϑ_0^* in Θ^*. □

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10463-023-00889-z.

References

Berkes, I., Horváth, L., Kokoszka, P. (2004). Testing for parameter constancy in GARCH(p, q) models. *Statistics & Probability Letters, 70*, 263–273.

Billingsley, P. (1968). *Convergence of probability measure*. New York: Wiley.

Bloomfield, P., Steiger, W. L. (1983). *Least absolute deviations: Theory, applications, and algorithms*. Boston: Birkhäuser.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327.

Bougerol, P., Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics, 52*, 115–127.

Campbell, S. D. (2006). A review of backtesting and backtesting procedures. *The Journal of Risk, 9*, 1–17.

Chen, J., Gupta, A. K. (2012). *Parametric statistical change point analysis with applications to genetics, medicine, and finance*. New York: Wiley.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review, 39*, 841–862.

Ciuperca, G. (2017). Real time change-point detection in a nonlinear quantile model. *Sequential Analysis, 36*, 87–110.

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

de Pooter, M., van Dijk, D. (2004). *Testing for changes in volatility in heteroskedastic time series—A further examination*. Econometric Institute Research Papers EI 2004-38, Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987–1007.

Engle, R., Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics, 22*, 367–381.

Fitzenberger, B., Koenker, R., Machado, J. (2013). *Economic applications of quantile regression*. Berlin, Heidelberg: Springer.

Francq, C., Zakoian, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli, 10*, 605–637.

Gombay, E. (2008). Change detection in autoregressive time series. *Journal of Multivariate Analysis, 99*, 51–464.

Hall, P., Heyde, C. C. (1980). *Martingale limit theory and its application*. San Diego: Academic Press.

Inclán, C., Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association, 89*, 913–923.

Kim, M., Lee, S. (2016). Nonlinear expectile regression with application to value-at-risk and expected shortfall estimation. *Computational Statistics & Data Analysis, 94*, 1–19.

Kim, S., Cho, S., Lee, S. (2000). On the cusum test for parameter changes in GARCH(1, 1) models. *Communications in Statistics-Theory and Methods, 29*, 445–462.

Kirch, C., Kamgaing, J. (2012). Testing for parameter stability in nonlinear autoregressive models. *Journal of Time Series Analysis, 33*, 365–385.

Koenker, R., Bassett, Jr., G. (1978). Regression quantiles. *Econometrica: Journal of the Econometric Society*, **46**, 33–50.
Koenker, R., Zhao, Q. (1996). Conditional quantile estimation and inference for ARCH models. *Econometric Theory, 12*, 793–813.
Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives, 3*, 73–84.
Lee, S. (2020). Location and scale-based CUSUM test with application to autoregressive models. *Journal of Statistical Computation and Simulation, 90*, 2309–2328.
Lee, S., Kim, C. K. (2022). Test for conditional quantile change in GARCH models. *Journal of the Korean Statistical Society, 51*, 480–499.
Lee, S., Noh, J. (2010). Value at risk forecasting based on quantile regression for GARCH models. *The Korean Journal of Applied Statistics, 23*, 669–681.
Lee, S., Noh, J. (2013). Quantile regression estimator for GARCH models. *Scandinavian Journal of Statistics, 40*, 2–20.
Lee, S., Ha, J., Na, O., Na, S. (2003). The CUSUM test for parameter change in time series models. *Scandinavian Journal of Statistics, 30*, 781–796.
Lee, S., Tokutsu, Y., Maekawa, K. (2004). The cusum test for parameter change in regression models with ARCH errors. *Journal of Japan Statistical Society, 34*, 173–188.
Meitz, M., Saikkonen, P. (2008). Ergodicity, mixing, and existence of moments of a class of Markov models with applications to GARCH and ACD models. *Econometric Theory, 24*, 1291–1320.
Noh, J., Lee, S. (2016). Quantile regression for location-scale time series models with conditional heteroscedasticity. *Scandinavian Journal of Statistics, 43*, 700–720.
Oh, H., Lee, S. (2018). On score vector-and residual-based CUSUM tests in ARMA-GARCH models. *Statistical Methods & Applications, 27*, 385–406.
Oh, H., Lee, S. (2019). Modified residual CUSUM test for location-scale time series models with heteroscedasticity. *Annals of Institute of Statistical Mathematics, 71*, 1059–1091.
Oka, T., Qu, Z. (2011). Estimating structural changes in regression quantiles. *Journal of Econometrics, 162*, 248–267.
Page, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika, 42*, 523–527.
Qu, Z. (2008). Testing for structural change in regression quantiles. *Journal of Econometrics, 146*, 170–184.
Straumann, D., Mikosch, T. (2006). Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics, 34*, 2449–2495.
Su, L., Xiao, Z. (2008). Testing for parameter stability in quantile regression models. *Statistics & Probability Letters, 78*, 2768–2775.
Weiss, A. (1991). Estimating nonlinear dynamic models using least absolute error estimation. *Econometric Theory, 7*, 46–68.
Xiao, Z., Koenker, R. (2009). Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models. *Journal of the American Statistical Association, 104*, 1696–1712.
Zhou, M., Wang, H. J., Tang, Y. (2015). Sequential change point detection in linear quantile regression models. *Statistics & Probability Letters, 100*, 98–103.