Annals of the Institute of Statistical Mathematics (2019) 71:1059–1091
https://doi.org/10.1007/s10463-018-0679-4

Modified residual CUSUM test for location-scale time series models with heteroscedasticity

Haejune Oh[1] · **Sangyeol Lee**[1]

Received: 28 June 2017 / Revised: 12 June 2018 / Published online: 23 July 2018

Abstract
This study considers the residual-based CUSUM test for location-scale time series models with heteroscedasticity. The estimates- and score vector-based CUSUM tests are widely used for detecting abrupt changes in time series models. However, their performance is often unsatisfactory with severe size distortions when the underlying model is complicated and the sample size is small. To circumvent this defect, the residual-based CUSUM test is suggested as an alternative. However, this test can only detect scale parameter changes and suffers severe power loss against location parameter changes. To remedy this, we introduce a modified residual-based CUSUM test and demonstrate its validity for both location and scale parameter changes. We conduct a simulation study and data analysis for illustration.

Keywords Location-scale time series models with heteroscedasticity · Parameter change test · CUSUM test · Residual-based test · Score vector-based test

1 Introduction

This study considers the residual-based CUSUM test for location-scale time series models with heteroscedasticity. Since Page (1955), the problem of testing for a parameter change has been an important issue in economics, engineering and medicine, and a multitude of articles have been published in various research areas. The change point problem has drawn much attention from many researchers in time series analysis, as

This work 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. 2018R1A2A2A05019433).

✉ Sangyeol Lee
sylee@stats.snu.ac.kr

Haejune Oh
haejune.oh@gmail.com

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

time series often exhibit structural changes owing to changes in policy and critical social events. It is widely appreciated that detecting a change point is crucial and ignoring it can lead to a false conclusion. The literature on the change point tests for time series models is quite extensive. The CUSUM test of Brown et al. (1975) was first applied by Inclán and Tiao (1994) to detect multiple changes of variance in independent samples. Since then, numerous studies have been conducted on the CUSUM test for GARCH-type models: See Kim et al. (2000) and Kokoszka and Leipus (1999) for earlier studies. Kulperger and Yu (2005) study high moment partial sum processes based on residuals and apply them to the residual CUSUM test for GARCH models. Berkes et al. (2004) and Gombay (2008) consider the score vector-based CUSUM test in GARCH and AR models. Lee and Na (2005) propose the CUSUM test based on conditional least squares estimators. Lee and Song (2008) and Lee and Oh (2016) also study the estimates-based CUSUM test in ARMA-GARCH and ACD models.

The conventional estimates-based CUSUM test is designed to compare the discrepancy among sequentially obtained estimators: See Lee et al. (2003). This estimates-based test performs well in general but suffers from severe size distortions and produces low powers on some occasions; particularly when the underlying model is complicated and has many unknown parameters, the parameters lie near the border of stationary domains, the sample size is relatively small and the error distribution is highly nonnormal: See Kang and Lee (2014) and Lee and Lee (2015). The residual-based CUSUM test has been proposed as a remedy, since it develops a more stable test owing to the removal of the dependency of time series: See Lee et al. (2004), Kang and Lee (2014) and Lee and Lee (2015). The advantages of the residual-based CUSUM test over the estimates-based CUSUM test are also advocated by de Pooter and van Dijk (2004). However, the residual-based CUSUM test for location-scale models often suffers from a severe power loss in detecting location parameter changes as seen in AR(1) models, since it only responds to scale parameter changes, where "location parameter' and "scale parameter," respectively, indicate the parameters in conditional mean and variance of location-scale models. To overcome this drawback, Oh and Lee (2017a) suggest using the score vector-based CUSUM test for ARMA-GARCH models. However, despite its own merits, the test still exhibits nontrivial size distortions in some situations.

To resolve this problem, we introduce a modified residual-based CUSUM test that can effectively detect both location and scale parameter changes. This test is much simpler to implement than the score vector-based CUSUM test because no extra steps are required to calculate the derivatives of location and scale components and the proposed test statistic is only two-dimensional, representing the location and scale components, whereas the score vector-based CUSUM test statistic is multi-dimensional, proportional to the number of model parameters. Moreover, it performs better in terms of stability and power than the estimates- and original residual-based CUSUM tests and the score vector-based CUSUM test, as seen on some occasions in our simulation study.

The organization of this paper is as follows. In Sect. 2, we introduce a modified residual-based CUSUM test. In Sect. 3, we perform a simulation study. Section 4 gives an example of real data analysis using Dow30 datasets. Section 5 provides concluding remarks, and Sect. 6 contains all the lemmas and proofs.

2 Residual-based CUSUM test

2.1 CUSUM test for location-scale models

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

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

where $g_t(\mu) = g(y_{t-1}, y_{t-2}, \ldots; \mu)$ and $h_t(\theta) = h(y_{t-1}, y_{t-2}, \ldots; \theta)$ for $\mu \in \Theta_1$, $\theta = (\mu^{\mathrm{T}}, \lambda^{\mathrm{T}})^{\mathrm{T}} \in \Theta = \Theta_1 \times \Theta_2 \subset \mathbf{R}^m$ with compact subsets $\Theta_1 \subset \mathbf{R}^{m_1}$ and $\Theta_2 \subset \mathbf{R}^{m_2}$; $g : \mathbf{R}^\infty \times \Theta_1 \to \mathbf{R}$ and $h : \mathbf{R}^\infty \times \Theta \to \mathbf{R}$ are known measurable functions; $\theta_0 = (\mu_0^{\mathrm{T}}, \lambda_0^{\mathrm{T}})^{\mathrm{T}}$ is the true parameter belonging to the interior of Θ; $\{\eta_t\}$ is a sequence of i.i.d. random variables with mean zero and unit variance. Model (1) includes a broad class of conditionally heteroscedastic time series models, covering invertible ARMA models and stationary GARCH models.

In what follows, we denote by $\mathcal{F}_t = \sigma(\eta_s : s \le t)$ the σ-field generated by $\{\eta_s : s \le t\}$ and assume the following conditions:

(M1) $\{y_t\}$ is $\{\mathcal{F}_t\}$-adapted, strictly stationary and ergodic with $Ey_t = 0$.
(M2) η_t is independent of $\mathcal{F}_s$ for $s < t$ and $\mathbb{E}\eta_t^4 < \infty$.

Given observations $y_1, \ldots, y_n$, our objective is to test the following hypotheses:

$$\begin{aligned} &H_0 : \theta = (\mu^{\mathrm{T}}, \lambda^{\mathrm{T}})^{\mathrm{T}} \text{ remains the same for the whole series v.s.} \\ &H_1 : \text{not } H_0. \end{aligned} \tag{2}$$

To conduct a test, we approximate $g_t(\mu)$ and η_t with their counterparts recursively computed with initial values: $\tilde{g}_t(\mu)$ and $\tilde{\eta}_t(\theta) = (y_t - \tilde{g}_t(\mu))/\sqrt{\tilde{h}_t(\theta)}$, $1 \le t \le n$. For instance, we can use $\tilde{g}_t(\mu) = g(y_{t-1}, y_{t-2}, \ldots, y_1, 0, \ldots; \mu)$ and $\tilde{h}_t(\theta) = h(y_{t-1}, y_{t-2}, \ldots, y_1, 0, \ldots; \theta)$. Then, we consider the CUSUM test based on $\{(\tilde{g}_t(\mu)\tilde{\eta}_t(\theta), \tilde{\eta}_t^2(\theta))\}$ with θ_0 replaced by its estimator $\hat{\theta}_n = (\hat{\mu}^{\mathrm{T}}, \hat{\lambda}^{\mathrm{T}})^{\mathrm{T}}$ as follows:

$$\hat{T}_n := \max_{1 \le k \le n} \frac{1}{n} \left(\sum_{t=1}^{k} \tilde{U}_t(\hat{\theta}_n) - \frac{k}{n} \sum_{t=1}^{n} \tilde{U}_t\left(\hat{\theta}_n\right) \right)^{\mathrm{T}} \hat{\Sigma}_n^{-1} \left(\sum_{t=1}^{k} \tilde{U}_t\left(\hat{\theta}_n\right) - \frac{k}{n} \sum_{t=1}^{n} \tilde{U}_t(\hat{\theta}_n) \right),$$

where $\tilde{U}_t(\theta) = (\tilde{g}_t(\mu)\tilde{\eta}_t(\theta), \tilde{\eta}_t^2(\theta))^{\mathrm{T}}$, $\hat{\Sigma}_n$ is a consistent estimator of $\Sigma_0 = \mathrm{Var}(U_1(\theta_0))$, and $U_1(\theta_0) = (g_1(\mu_0)\eta_1, \eta_1^2)^{\mathrm{T}}$. Note that $\{\tilde{g}_t(\hat{\mu})\tilde{\eta}_t(\hat{\theta}_n)\}$ is newly introduced to enhance the capability to detect the change in location parameter μ, whereas $\{\tilde{\eta}_t^2(\hat{\theta}_n)\}$ remains to detect the change of scale parameter λ. We do this because under the null of no changes, the mean of $g_t(\mu)\eta_t$ remains constant, viz. zero, while $\tilde{g}_t(\hat{\mu})\tilde{\eta}_t(\hat{\theta}_n)$ can detect a change in μ when used in the construction of the CUSUM test. In fact, the behind reasoning is essentially the same as the case of the scale parameter change based on $\{\tilde{\eta}_t^2(\hat{\theta}_n)\}$ (cf. Lee et al. 2004).

This test merits not to require the calculation of derivatives of g_t and h_t and escalates the efficacy of test compared with the score vector-based CUSUM test in application to more complicated time series models with many unknown parameters.

We impose some regularity conditions, wherein $L^2(\mathbf{R}^p)$, $p \geq 1$ and $L^2(\mathbf{R}^\infty)$, respectively, denote the class of all random vectors $X = (X_1, \ldots, X_p)$ and sequences $X = (X_1, X_2, \ldots)$ with $\sup_i \mathbb{E}X_i^2 < \infty$, and $0 < \rho < 1$ is a generic constant.

(A1) $h(\cdot;\theta)$ is continuous in $\theta \in \Theta$; for any $X, X' \in L^2(\mathbf{R}^\infty)$, $Y \in L^2(\mathbf{R}^{l-1})$ and integrable random variable $V_h(\cdot)$, $\sup_{\theta\in\Theta} |h(Y, X; \theta) - h(Y, X'; \theta)| \leq V_h(Y, X, X') \cdot \rho^l$ a.s.

(A2) $g(\cdot;\mu)$ is continuous in $\mu \in \Theta_1$; for any $X, X' \in L^2(\mathbf{R}^\infty)$, $Y \in L^2(\mathbf{R}^{l-1})$ and integrable random variable $V_g(\cdot)$, $\sup_{\theta\in\Theta} |g(Y, X; \mu) - g(Y, X'; \mu)| \leq V_g(Y, X, X') \cdot \rho^l$ a.s.

(A3) The function h is bounded below from 0, that is, $h \geq \underline{h} = \inf_{(X,\theta)\in R^\infty\times\Theta} h(X;\theta) > 0$.

(A4) $h(\cdot;\theta)$ is continuously differentiable with respect to θ on Θ; for any $X, X' \in L^2(\mathbf{R}^\infty)$, $Y \in L^2(\mathbf{R}^{l-1})$ and integrable random variable $V_{dh}(\cdot)$, $\sup_{\theta\in\Theta} \|\frac{\partial h(Y,X;\theta)}{\partial\theta} - \frac{\partial h(Y,X';\theta)}{\partial\theta}\| \leq V_{dh}(Y, X, X') \cdot \rho^l$ a.s.

(A5) $g(\cdot;\mu)$ is continuously differentiable with respect to μ on Θ_1; for any $X, X' \in L^2(\mathbf{R}^\infty)$, $Y \in L^2(\mathbf{R}^{l-1})$ and integrable random variable $V_{dg}(\cdot)$, $\sup_{\theta\in\Theta} \|\frac{\partial g(Y,X;\mu)}{\partial\mu} - \frac{\partial g(Y,X';\mu)}{\partial\mu}\| \leq V_{dg}(Y, X, X') \cdot \rho^l$ a.s.

(A6) The following moment conditions hold:

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

(ii) $\mathbb{E}\big(\log^+ \sup_{\theta\in\Theta} |h_t(\theta)|\big) < \infty$, $\mathbb{E}\big(\sup_{\theta\in\Theta} \frac{1}{h_t(\theta)} \|\frac{\partial h_t(\theta)}{\partial\theta}\|\big)^2 < \infty$.

(A7) Under the null, the estimator $\hat\theta_n = (\hat\mu_n^{\mathrm{T}}, \hat\lambda_n^{\mathrm{T}})^{\mathrm{T}}$ of $\theta_0 = (\mu_0^{\mathrm{T}}, \lambda_0^{\mathrm{T}})^{\mathrm{T}}$ satisfies (i) $\sqrt{n}(\hat\mu_n - \mu_0) = O_{\mathrm{P}}(1)$ and (ii) $\sqrt{n}(\hat\lambda_n - \lambda_0) = O_{\mathrm{P}}(1)$.

For example, the above conditions hold for the AR(1)-GARCH(1,1) model:

$$
\begin{aligned}
y_t &= \phi y_{t-1} + \epsilon_t,\\
\epsilon_t &= \sqrt{h_t}\eta_t, \quad h_t = \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1},
\end{aligned}
$$

where $\{\eta_t\}$ is a sequence of i.i.d. random variables with $\mathbb{E}\eta_t = 0$, $\mathbb{E}\eta_t^2 = 1$ and $\mathbb{E}\eta_t^4 < \infty$. We assume $|\phi| < 1$, $\phi \neq 0$, $0 < \alpha, \beta < 1$, and $\alpha^2\mathbb{E}\eta_t^4 + 2\alpha\beta + \beta^2 < 1$. Then, $\{y_t\}$ is strictly stationary and ergodic (see Sect. 2.3.2) with $\mathbb{E}y_t^4 < \infty$. In this case, $\mu = \phi$, $\lambda = (\omega, \alpha, \beta)^{\mathrm{T}}$, $\theta = (\phi, \omega, \alpha, \beta)^{\mathrm{T}}$, $g_t(\mu) = \phi y_{t-1}$ and $h_t(\theta) = \frac{\omega}{1-\beta} + \alpha\sum_{k=1}^\infty \beta^{k-1}(y_{t-k} - \phi y_{t-k-1})^2$. We assume that for some positive numbers $\underline{\omega}, \bar\omega, \underline{\alpha}, \bar\alpha, \underline{\beta}, \bar\beta$, $\underline{\omega} \leq \omega \leq \bar\omega$, $\underline{\alpha} \leq \alpha \leq \bar\alpha$, $\underline{\beta} \leq \beta \leq \bar\beta$ and $\alpha + \beta \leq \rho_0$ with $0 < \rho_0 < 1$. Then, parameter space Θ is a compact subset of $\mathbf{R}^4$, consisting of members satisfying these conditions. Let $\theta_0 = (\mu_0, \lambda_0)^{\mathrm{T}}$ be an interior of Θ, where $\mu_0 = \phi_0$ and $\lambda_0 = (\omega_0, \alpha_0, \beta_0)$. For any $X, X' \in L^2(\mathbf{R}^\infty)$, $Y \in L^2(\mathbf{R}^{l-1})$, we have $h(Y, X; \theta) = \frac{\omega}{1-\beta} + \alpha\sum_{k=1}^{l-2} \beta^{k-1}(Y_k - \phi Y_{k+1})^2 + \alpha\beta^{l-2}(Y_{l-1} - \phi X_1)^2 + \alpha\sum_{k=1}^\infty \beta^{l-1+k}(X_k - \phi X_{k+1})^2$. Notice that (A1) holds since $\sup_{\theta\in\Theta} |h(Y, X; \theta) - h(Y, X'; \theta)| = \sup_{\theta\in\Theta} \Big|\beta^l \frac{\alpha}{\beta^2}\Big((Y_{l-1} - \phi X_1)^2 - (Y_{l-1} - \phi X_1')^2 + \sum_{k=1}^\infty\{(X_k - \phi X_{k+1})^2 - (X_k' - \phi X_{k+1}')^2\}\Big)\Big| \leq \rho_0^l \cdot V_h(Y, X, X')$ a.s. It is also easy to check that (A2)–(A5) hold. Since Gaussian QMLE $\hat{\theta}_n$ in AR(1)-GARCH(1,1) models is strongly consistent and asymptotic normal (see Sect. 2.3.2), (A7) holds.

We then obtain the following result.

Theorem 1 *Assume that* (M1), (M2) *and* (A1)–(A7) *hold. Then, under* H_0*, as* $n \to \infty$,

$$\hat{T}_n \xrightarrow{w} \sup_{0\le s\le 1} \left\| W_2^{\circ}(s) \right\|^2, \tag{3}$$

where $W_d^{\circ}(\cdot)$ *denotes a* d*-dimensional Brownian bridge (with independent components).*

We reject H_0 if $\hat{T}_n \ge C_\alpha$ at the nominal level α, where C_α is the $100(1-\alpha)$ quantile value of $\sup_{0\le s\le 1} \|W_d^{\circ}(s)\|^2$ with $d = 2$. The critical values are provided in Table 1 in Kiefer (1959) for $d = 1, \ldots, 5$ and Table 1 in Lee et al. (2003) for $d = 1, \ldots, 10$. In implementation, as an estimate of Σ_0, we can employ

$$\hat{\Sigma}_n = \begin{pmatrix} \frac{1}{n}\sum_{t=1}^{n} \tilde{g}_t^2(\hat{\mu}_n)\tilde{\eta}_t^2(\hat{\theta}_n) & \frac{1}{n}\sum_{t=1}^{n} \tilde{g}_t(\hat{\mu}_n)\tilde{\eta}_t^3(\hat{\theta}_n) \\ \frac{1}{n}\sum_{t=1}^{n} \tilde{g}_t(\hat{\mu}_n)\tilde{\eta}_t^3(\hat{\theta}_n) & \frac{1}{n}\sum_{t=1}^{n} \tilde{\eta}_t^4(\hat{\theta}_n) - 1 \end{pmatrix}. \tag{4}$$

If one wishes to test for a parameter change in the following model:

$$y_t = g_t(\mu_0) + \sigma_0\eta_t \quad \text{for} \quad t \in Z \tag{5}$$

with $\theta_0 = (\mu_0^{\mathrm{T}}, \sigma_0)^{\mathrm{T}}$, one can construct the CUSUM test based on $\{\tilde{g}_t(\hat{\mu}_n)\tilde{\eta}_t(\hat{\theta}_n)\}$ and obtain the following.

Corollary 1 *Assume that* (M1), (M2), (A2), (A5), (A6)(i) *and* (A7) *hold. Then, under* H_0*, as* $n \to \infty$,

$$\begin{aligned} \hat{T}_n^L &:= \frac{1}{\sqrt{n}\hat{\kappa}_n} \max_{1\le k\le n} \left| \sum_{t=1}^{k} \tilde{g}_t(\hat{\mu}_n)\tilde{\eta}_t\left(\hat{\theta}_n\right) - \frac{k}{n}\sum_{t=1}^{n} \tilde{g}_t\left(\hat{\mu}_n\right)\tilde{\eta}_t\left(\hat{\theta}_n\right) \right| \\ &\xrightarrow{w} \sup_{0\le s\le 1} |W_1^{\circ}(s)|, \end{aligned} \tag{6}$$

where $\hat{\kappa}_n^2 = \frac{1}{n}\sum_{t=1}^{n} \tilde{g}_t^2(\hat{\mu}_n)\tilde{\eta}_t^2(\hat{\theta}_n)$ *and* L *stands for "location."*

On the other hand, for the scale parameter change test in the following model:

$$y_t = c_0 + \sqrt{h_t(\theta_0)}\eta_t, \quad t \in Z \tag{7}$$

with $\theta_0 = (c_0, \lambda_0^{\mathrm{T}})^{\mathrm{T}}$, we can use the CUSUM test based on the $\{\tilde{\eta}_t^2(\hat{\theta}_n)\}$, and obtain the following.

Corollary 2 *Assume that* (M1), (M2), (A1), (A3), (A4), (A6)(ii) *and* (A7) *hold. Then, under* H_0*, as* $n \to \infty$,

$$\hat{T}_n^{\mathrm{Res}} := \frac{1}{\sqrt{n}\hat{\nu}_n} \max_{1\le k\le n} \left| \sum_{t=1}^{k} \tilde{\eta}_t^2(\hat{\theta}_n) - \frac{k}{n}\sum_{t=1}^{n} \tilde{\eta}_t^2(\hat{\theta}_n) \right| \xrightarrow{w} \sup_{0\le s\le 1} |W_1^{\circ}(s)|,$$

where $\hat{\nu}_n^2 = \frac{1}{n}\sum_{t=1}^{n} \tilde{\eta}_t^4(\hat{\theta}_n) - 1$ *and Res stands for "residuals."*

Table 1 Empirical sizes and powers for the ARMA(1,1)-GARCH(1,1) model with $N(0, 1)$ errors

$(\phi, \psi, \omega, \alpha, \beta)$		$n = 500$		$n = 800$	
Nominal level		0.05	0.10	0.05	0.10
	$\hat{T}_n$	**0.046**	**0.084**	**0.052**	**0.092**
Size	$\hat{T}_n^L$	0.050	0.090	0.056	0.098
(0.5, 0.1, 0.2, 0.4, 0.2)	$\hat{T}_n^{\text{Res}}$	0.036	0.082	0.040	0.094
	$\hat{T}_n^S$	0.042	0.100	0.064	0.116
	$\hat{T}_n^E$	0.278	0.346	0.178	0.234
	$\hat{T}_n$	**0.414**	**0.532**	**0.620**	**0.726**
ϕ	$\hat{T}_n^L$	0.496	0.582	0.702	0.784
$0.5 \longrightarrow 0.7$	$\hat{T}_n^{\text{Res}}$	0.030	0.076	0.040	0.094
	$\hat{T}_n^S$	0.400	0.506	0.636	0.742
	$\hat{T}_n^E$	0.722	0.780	0.852	0.892
	$\hat{T}_n$	**0.234**	**0.354**	**0.398**	**0.510**
ψ	$\hat{T}_n^L$	0.358	0.432	0.480	0.564
$0.1 \longrightarrow 0.3$	$\hat{T}_n^{\text{Res}}$	0.030	0.062	0.046	0.094
	$\hat{T}_n^S$	0.284	0.378	0.424	0.530
	$\hat{T}_n^E$	0.390	0.484	0.436	0.554
	$\hat{T}_n$	**0.724**	**0.816**	**0.948**	**0.968**
ω	$\hat{T}_n^L$	0.046	0.088	0.062	0.112
$0.2 \longrightarrow 0.4$	$\hat{T}_n^{\text{Res}}$	0.830	0.896	0.976	0.992
	$\hat{T}_n^S$	0.606	0.750	0.918	0.956
	$\hat{T}_n^E$	0.768	0.844	0.928	0.968
	$\hat{T}_n$	**0.128**	**0.206**	**0.212**	**0.316**
α	$\hat{T}_n^L$	0.062	0.102	0.070	0.102
$0.4 \longrightarrow 0.6$	$\hat{T}_n^{\text{Res}}$	0.170	0.258	0.276	0.354
	$\hat{T}_n^S$	0.106	0.180	0.156	0.260
	$\hat{T}_n^E$	0.500	0.584	0.534	0.620
	$\hat{T}_n$	**0.374**	**0.522**	**0.650**	**0.784**
β	$\hat{T}_n^L$	0.058	0.108	0.066	0.104
$0.2 \longrightarrow 0.4$	$\hat{T}_n^{\text{Res}}$	0.526	0.638	0.766	0.860
	$\hat{T}_n^S$	0.252	0.410	0.494	0.670
	$\hat{T}_n^E$	0.736	0.786	0.876	0.904
ψ	$\hat{T}_n$	**0.564**	**0.710**	**0.864**	**0.918**
$0.1 \longrightarrow 0.3$	$\hat{T}_n^L$	0.310	0.424	0.472	0.572
β	$\hat{T}_n^{\text{Res}}$	0.470	0.604	0.766	0.860
$0.2 \longrightarrow 0.4$	$\hat{T}_n^S$	0.582	0.690	0.818	0.898
	$\hat{T}_n^E$	0.846	0.882	0.950	0.974

2.2 Consistency of the CUSUM test

We investigate the consistency of the modified CUSUM test under the alternative of one change point. We introduce the two independent time series $\{y_{1,t}\}$ and $\{y_{2,t}\}$ generated from model (1) with $\theta_1 = (\mu_1^T, \lambda_1^T)^T \neq \theta_2 = (\mu_2^T, \lambda_2^T)^T$ in Θ, respectively. We then consider $y_1, \ldots, y_n$ as follows:

$$y_t = \begin{cases} y_{1,t} & t \leq k_0, \\ y_{2,t} & t > k_0. \end{cases}$$

We assume that the following conditions hold:

(M3) $\{y_{1,t}\}$ and $\{y_{2,t}\}$ are $\{\mathcal{F}_t\}$-adapted, strictly stationary and ergodic.
(M4) The change point fulfills $k_0 = [n\tau_0]$ for some $0 < \tau_0 < 1$.

For notational convenience, we express $g(y_{i,t-1}, y_{i,t-2}, \ldots; \mu)$ as $g_{i,t}(\mu)$ and $\frac{y_{i,t} - g(y_{i,t-1}, y_{i,t-2}, \ldots; \mu)}{\sqrt{h(y_{i,t-1}, y_{i,t-2}, \ldots; \theta)}}$ as $\eta_{i,t}(\theta)$ for $i = 1, 2$ and assume the following conditions:

(C1) There exist $\tilde{\mu}_0 \in \Theta_1$ and positive constants $E_{1,\lambda}$ for $\lambda = \lambda_0$ and $\tilde{\lambda}_0$ defined in (C2), such that

(i) $\sqrt{n}(\hat{\mu}_n - \tilde{\mu}_0) = O_P(1)$,
(ii) $E_{1,\lambda} = |\mathbb{E}g_{1,1}(\tilde{\mu}_0)\eta_{1,1}(\tilde{\mu}_0, \lambda) - \mathbb{E}g_{2,1}(\tilde{\mu}_0)\eta_{2,1}(\tilde{\mu}_0, \lambda)|$.

(C2) There exist $\tilde{\lambda}_0 \in \Theta_2$ and constants $E_{2,\mu}$ for $\mu = \mu_0$ and $\tilde{\mu}_0$, such that

(i) $\sqrt{n}(\hat{\lambda}_n - \tilde{\lambda}_0) = O_P(1)$,
(ii) $E_{2,\mu} = |\mathbb{E}\eta_{1,1}^2(\mu, \tilde{\lambda}_0) - \mathbb{E}\eta_{2,1}^2(\mu, \tilde{\lambda}_0)|$.

Instead of (M3), one may consider an alternative setting in Berkes et al. (2004), where $\{y_{2t}\}$ is a stationary process not dependent on the past values of $\{y_{1t}\}$ before the change point. In our setup, we regard initial values of $\{y_{2t}\}$, which are actually from $\{y_{1t}\}$, as fixed real numbers. Concerning (C1) and (C2), we refer to Gombay (2008), Kirch and Kamgaing (2012) and Franke et al. (2012) who study the asymptotic properties of the quasi-maximum likelihood (QML) estimates of linear AR(p) models, the nonlinear least squares (NLLS) estimates for nonlinear AR models and conditional least squares (CLS) estimates for the general Poisson AR models under the alternative of one change point. Although no results are available for other cases in the literature, they are not unrealistic assumptions and quite likely to hold in many situations.

As seen in Theorems 2 and 3, conditions (C1)(ii) and (C2)(ii) play an important role. In fact, if $\mathbb{E}g_{i,t}(\tilde{\mu})\eta_{i,t}(\tilde{\theta}), i = 1, 2$, are distinct (nonzero) constants, we have $E_{1,\lambda} > 0$, and further, if $\mathbb{E}\eta_{i,t}^2(\tilde{\theta}), i = 1, 2$, are distinct, we have $E_{2,\mu} > 0$. These are easy to conjecture conceptually, but in general, are not easy to show analytically. We provide an empirical evidence in our simulation study, wherein $E_{1,\lambda}$ and $E_{2,\mu}$ are calculated for TAR(1)-GARCH(1,1) and Logistic STAR(1)-STGARCH(1,1) models: See Tables 8 and 9.

Theorem 2 *Assume that* (M2)–(M4) *and* (A1)–(A6) *and* (C1), (C2)(i) *hold. Then, as* $n \to \infty$,

$$\hat{T}_n^L \xrightarrow{P} \infty \quad and \quad \frac{\hat{k}_n^L}{n} \xrightarrow{P} \tau_0,$$

where $\hat{k}_n^L = \text{argmax}\{|\hat{T}_{n,k}^L| : 1 < k < n\}$ *and* $\hat{T}_{n,k}^L := \sum_{t=1}^k \tilde{g}_t(\hat{\mu}_n)\tilde{\eta}_t(\hat{\theta}_n) - \frac{k}{n}\sum_{t=1}^n \tilde{g}_t(\hat{\mu}_n)\tilde{\eta}_t(\hat{\theta}_n)$.

Theorem 3 *Assume that* (M2)–(M4) *and* (A1)–(A6) *and* (C1)(i), (C2) *hold. Then, as* $n \to \infty$,

$$\hat{T}_n^{\text{Res}} \xrightarrow{P} \infty \quad and \quad \frac{\hat{k}_n^{\text{Res}}}{n} \xrightarrow{P} \tau_0,$$

where $\hat{k}_n^{\text{Res}} = \text{argmax}\{|\hat{T}_{n,k}^{\text{Res}}| : 1 < k < n\}$ *and* $\hat{T}_{n,k}^{\text{Res}} := \sum_{t=1}^k \tilde{\eta}_t^2(\hat{\theta}_n) - \frac{k}{n}\sum_{t=1}^n \tilde{\eta}_t^2(\hat{\theta}_n)$.

2.3 Some applications

In this section, we consider AR and ARMA-GARCH models as examples. We verify that the AR(1) model satisfies the regularity conditions in Sects. 2.1 and 2.2 under mild conditions. In ARMA-GARCH models, more concrete conditions are discussed to check the regularity conditions.

2.3.1 AR(1) models

Consider the AR(1) model:

$$y_t = \phi_0 y_{t-1} + \sigma\eta_t,$$

where $\{\eta_t\}$ are i.i.d random variables with $\mathbb{E}\eta_t = 0$, $\mathbb{E}\eta_t^2 = 1$, and σ is a positive constant. The parameter ϕ belongs to compact subset of $\mathbf{R}$, containing the true parameter ϕ_0 as its interior point. In this case, the conditional mean $g_t(\phi)$ is ϕy_{t-1}. Moreover, the CLS(or QML) estimator of ϕ_0 is given by

$$\hat{\phi}_n = \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} = \frac{\sum_{t=1}^n \tilde{y}_t \tilde{y}_{t-1}}{\sum_{t=1}^n \tilde{y}_{t-1}^2},$$

where $\tilde{y}_0$ is 0 and $\tilde{y}_t = y_t$ for $1 \le t \le n$. Suppose that

(a1) $|\phi_0| < 1$,
(a2) $\eta_t^4 < \infty$.

Then, we can easily check that conditions (M1), (A2), (A5), (A6)(i) and (A7) are satisfied. Hence, when there is no parameter change in ϕ, (6) holds.

Next, assume that

(b1) $|\phi_1| < 1, |\phi_2| < 1$ and $\phi_1 \neq \phi_2$.
(b2) There is a parameter change from ϕ_1 to ϕ_2 at $k_0 = [n\tau_0]$ for some $0 < \tau_0 < 1$.

Then, (b1) implies (M3) and we can easily see that

$$\hat{\phi}_n \to \tilde{\phi}_0 := \nu\phi_1 + (1-\nu)\phi_2 \tag{8}$$

a.s. with $\nu = \frac{\tau_0/(1-\phi_1^2)}{\tau_0/(1-\phi_1^2)+(1-\tau_0)/(1-\phi_2^2)}$. Moreover, $\sqrt{n}(\hat{\phi}_n - \tilde{\phi}_0)$ asymptotically follows a normal distribution: Note that $\sqrt{n}(\hat{\phi}_{1n} - \phi_1)$ and $\sqrt{n}(\hat{\phi}_{2n} - \phi_2)$ asymptotically follow a normal distribution with $\hat{\phi}_{1n} = \frac{\sum_{t=2}^{k_0} y_t y_{t-1}}{\sum_{t=2}^{k_0} y_{t-1}^2}$ and $\hat{\phi}_{2n} = \frac{\sum_{t=k_0+1}^{n} y_t y_{t-1}}{\sum_{t=k_0+1}^{n} y_{t-1}^2}$. This implies (C1)(i).

Finally, we show that (C1)(ii) holds with $\tilde{\phi}_0$. Note that

$$\mathbb{E}g_1(\phi)\eta_1(\phi) = \mathbb{E}[\phi y_1(y_1 - \phi y_0)],$$

and thus,

$$\begin{aligned} &\left|\mathbb{E}g_{1,1}\left(\tilde{\phi}_0\right)\eta_{1,1}\left(\tilde{\phi}_0\right) - \mathbb{E}g_{2,1}\left(\tilde{\phi}_0\right)\eta_{2,1}\left(\tilde{\phi}_0\right)\right| \\ &= \left|\sigma^2\tilde{\phi}_0\left[\phi_1/\left(1-\phi_1^2\right) - \phi_2/\left(1-\phi_2^2\right) - \tilde{\phi}_0\left\{1/\left(1-\phi_1^2\right) + 1/\left(1-\phi_2^2\right)\right\}\right]\right|. \end{aligned}$$

If $|\phi_1| < 1, |\phi_2| < 1$ and $\phi_1 \neq \phi_2$,

$$\left[\phi_1/\left(1-\phi_1^2\right) - \phi_2/\left(1-\phi_2^2\right) - \tilde{\phi}_0\left\{1/\left(1-\phi_1^2\right) + 1/\left(1-\phi_2^2\right)\right\}\right] \neq 0.$$

This indicates that $|\mathbb{E}g_{1,1}(\tilde{\phi}_0)\eta_{1,1}(\tilde{\phi}_0) - \mathbb{E}g_{2,1}(\tilde{\phi}_0)\eta_{2,1}(\tilde{\phi}_0)| \neq 0$ unless otherwise $\tilde{\phi}_0 = 0$, which will occur with very low possibilities.

2.3.2 ARMA-GARCH models

Consider the ARMA(p,q) models with GARCH(r,s) innovations:

$$\begin{aligned} y_t &= \sum_{i=1}^{p}\phi_{0i}y_{t-1} + e_t + \sum_{j=1}^{q}\psi_{0j}e_{t-j}, \\ e_t &= \sqrt{h_t}\eta_t, \quad h_t = \omega_0 + \sum_{i=1}^{s}\alpha_{0i}e_{t-i}^2 + \sum_{j=1}^{r}\beta_{0j}h_{t-j}, \end{aligned} \tag{9}$$

where $\omega_0 > 0$, $\alpha_{0i} \geq 0$, $i = 1, \ldots, s$, $\beta_{0j} \geq 0$, $j = 1, \ldots, q$, and $\{\eta_t\}$ is a sequence of i.i.d. r.v.'s such that $\mathbb{E}\eta_t = 0$ and $\mathbb{E}\eta_t^2 = 1$. We set $\theta = (\mu^T, \lambda^T)^T \subset \Theta \in \mathbf{R}^{p+q+1}\times]0, +\infty[\times[0, \infty[^{s+r}$, where $\mu = (\phi_1, \ldots, \phi_p, \psi_1, \ldots, \psi_q)^T$ and $\lambda = (\omega, \alpha_1, \ldots, \alpha_s, \beta_1, \ldots, \beta_r)^T$ and denote the true parameter by $\theta_0 = (\mu_0^T, \lambda_0^T)^T$.

We assume the following to guarantee that a strictly stationary solution $\{e_t\}$ exists with a finite fourth moment (see Chen and An 1998):

(c0) The spectral radius of $\mathbb{E}(A_{0t} \otimes A_{0t}) < 1$, where $\otimes$ denotes Kronecker product and

$$A_{0t} := \begin{pmatrix} \alpha_{01}\eta_t^2 & \cdots & \alpha_{0s}\eta_t^2 & \beta_{01}\eta_t^2 & \cdots & \beta_{0r}\eta_t^2 \\ & & 0 & 0 & \cdots & 0 \\ & \mathbf{I}_{s-1} & \vdots & \vdots & \ddots & \vdots \\ & & 0 & 0 & \cdots & 0 \\ \alpha_{01} & \cdots & \alpha_{0s} & \beta_{01} & \cdots & \beta_{0r} \\ 0 & \cdots & 0 & & & 0 \\ \vdots & \ddots & \vdots & & \mathbf{I}_{r-1} & \vdots \\ 0 & \cdots & 0 & & & 0 \end{pmatrix}.$$

Further, we assume

(c1) For all $\theta \in \Theta$, $\Phi_\mu(z)\Psi_\mu(z) = 0$ implies $|z| > 1$, where $\Phi_\mu(z) = 1 - \sum_{i=1}^p \phi_i z^i$ and $\Psi_\mu(z) = 1 - \sum_{j=1}^q \psi_j z^j$.
(c2) For all $\theta \in \Theta$, $\sum_{j=1}^r \beta_j < 1$.
(c3) θ_0 lies in the interior of the compact set Θ.
(c4) If $r > 0$, $\mathcal{A}_{\lambda_0}(z)$ and $\mathcal{B}_{\lambda_0}(z)$ have no common roots, $\mathcal{A}_{\lambda_0}(1) \neq 0$, and $\alpha_{0s} + \beta_{0r} \neq 0$, where $\mathcal{A}_\lambda(z) = \sum_{i=1}^s \alpha_i z^i$ and $\mathcal{B}_\lambda(z) = 1 - \sum_{j=1}^r \beta_j z^j$.
(c5) $\Phi_{\mu_0}(z)$ and $\Psi_{\mu_0}(z)$ have no common roots, $a_{0p} \neq 0$, or $b_{0q} \neq 0$.
(c6) There exists no set Λ of cardinality 2 such that $\mathrm{P}(\eta_t \in \Lambda) = 1$.

(c0) implies $\mathbb{E}e_t^4 < \infty$ and owing to (c0) and (c1), the solution $\{y_t\}$ of (9) is strictly stationary and ergodic. Thus, (M1) and (M2) hold. Let $\epsilon_t = \epsilon_t(\mu) = \Psi_\mu^{-1}(\mathbb{B})\Phi_\mu(\mathbb{B})y_t$, where $\mathbb{B}$ denotes the lag operator, and let $l_t = l_t(\theta) = \epsilon_t^2/\sigma_t^2 + \log\sigma_t^2$, where $\sigma_t^2 = \sigma_t^2(\theta)$ is the ergodic strictly stationary solution of

$$\sigma_t^2 = \omega + \sum_{i=1}^s \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2. \tag{10}$$

Note that $e_t = \epsilon_t(\mu_0) = \Psi_{\mu_0}^{-1}(\mathbb{B})\Phi_{\mu_0}(\mathbb{B})y_t$ and $h_t = \sigma_t^2(\theta_0) = \mathcal{B}_{\lambda_0}^{-1}(\mathbb{B})(\omega + \mathcal{A}_{\lambda_0}(\mathbb{B})e_t)$. Hence, model (9) admits the autoregressive representation in (1). Moreover, we can show that (A1)–(A6) hold by the arguments similar to those in the proofs of Theorems 3.1 and 3.2 of Francq and Zakoïan (2004).

To test the hypotheses in (2), we estimate θ_0 based on the Gaussian QMLE, defined as follows:

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmin}}\, \tilde{L}_n(\theta), \tag{11}$$

where $\tilde{L}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\tilde{l}_t(\theta)$, $\tilde{l}_t(\theta) = \tilde{\epsilon}_t^2(\mu)/\tilde{\sigma}_t^2(\theta) + \log\tilde{\sigma}_t^2(\theta)$ and $\tilde{\epsilon}_t(\mu)$, $\tilde{\sigma}_t^2(\theta)$ are defined recursively by

$$\tilde{\epsilon}_t = \tilde{\epsilon}_t(\mu) = y_t - \sum_{i=1}^{p}\phi_i\tilde{\sigma}\, y_{t-i} - \sum_{j=1}^{q}\psi_j\tilde{\epsilon}_{t-j},$$
$$\tilde{\sigma}_t^2 = \tilde{\sigma}_t^2(\theta) = \omega + \sum_{i=1}^{s}\alpha_i\tilde{\epsilon}_{t-i}^2 + \sum_{j=1}^{r}\beta_j\tilde{\sigma}_{t-j}^2,$$

and initial values are properly given as in Francq and Zakoïan (2004).

Francq and Zakoïan (2004) show that, under (c0)–(c6), $\hat{\theta}_n \to \theta_0$ a.s., and further, $\sqrt{n}(\hat{\theta}_n - \theta_0)$ is asymptotically distributed as $N(0, \Sigma)$, where $\Sigma = \mathcal{J}^{-1}\mathcal{I}\mathcal{J}^{-1}$ with finite and positive definite matrices:

$$\mathcal{I} = \mathbb{E}\left(\frac{\partial l_t(\theta_0)}{\partial\theta}\frac{\partial l_t(\theta_0)}{\partial\theta^{\mathrm{T}}}\right) \text{ and } \mathcal{J} = \mathbb{E}\left(\frac{\partial^2 l_t(\theta_0)}{\partial\theta\partial\theta^{\mathrm{T}}}\right),$$

which implies (A7).

Next, assume that

(d1) $\theta_1 = (\mu_1^{\mathrm{T}}, \lambda_1^{\mathrm{T}})^{\mathrm{T}}$ and $\theta_2 = (\mu_2^{\mathrm{T}}, \lambda_2^{\mathrm{T}})^{\mathrm{T}}$ satisfy the conditions (c3)–(c5), respectively, and $\theta_1 \neq \theta_2$.

(d2) There is a parameter change from θ_1 to θ_2 at $k_0 = [n\tau_0]$ for some $0 < \tau_0 < 1$.

Then, we redefine $l_t(\theta)$ by

$$l_t(\theta) = \begin{cases} l_{1,t}(\theta) = \log\sigma_{1,t}^2(\theta) + \frac{\epsilon_{1,t}^2(\mu)}{\sigma_{1,t}^2(\theta)} & t \le k_0, \\ l_{2,t}(\theta) = \log\sigma_{2,t}^2(\theta) + \frac{\epsilon_{2,t}^2(\mu)}{\sigma_{2,t}^2(\theta)} & t > k_0, \end{cases} \tag{12}$$

where $\epsilon_{i,t}(\mu)$ and $\sigma_{i,t}(\theta)$ are constructed by $y_{i,t}$ and $y_{i,t}$ is generated by θ_i, $i = 1, 2$. Then, $l_{i,t}(\theta)$ are strictly stationary and ergodic for any $\theta \in \Theta$. We can write $L_n(\theta) = \frac{1}{n}\big(\sum_{t=1}^{k_0} l_{1,t}(\theta) + \sum_{t=k_0+1}^{n} l_{2,t}(\theta)\big)$, and further,

$$\tilde{L}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\tilde{l}_t(\theta) = \frac{1}{n}\left(\sum_{t=1}^{k_0}\tilde{l}_{1,t}(\theta) + \sum_{t=k_0+1}^{n}\tilde{l}_{2,t}(\theta)\right), \tag{13}$$

where $\tilde{l}_{i,t}(\theta) = \log\tilde{h}_{i,t}(\theta) + \frac{\tilde{\epsilon}_{i,t}^2(\mu)}{\tilde{h}_{i,t}(\theta)}$ for $i = 1, 2$. In fact, $\tilde{h}_{i,t}(\theta)$ and $\tilde{\epsilon}_{i,t}(\theta)$ for $t \le k_0$ are functions of $y_{0,t}, \ldots, y_{0,1}, C$, and thus, we can write $\tilde{l}_t(\theta) = \tilde{l}_{1,t}(\theta) = l(y_{1,t}, y_{1,t-1}, \ldots, C; \theta)$. Similarly, for $t > k_0$, we write $\tilde{l}_t(\theta) = \tilde{l}_{2,t}(\theta) = l(y_{2,t}, \ldots, y_{2,k_0+1}, y_{1,k_0}, \ldots, y_{1,1}, C; \theta)$ with some initial values $y_{1,k_0}, \ldots, y_{1,1}, C$.

Define

$$\theta_{2-\tau} = \left(\mu_{2-\tau}^{\mathrm{T}}, \lambda_{2-\tau}^{\mathrm{T}}\right)^{\mathrm{T}} := \underset{\theta\in\Theta}{\operatorname{argmin}}\, E_\tau(\theta) \tag{14}$$

where $E_\tau(\theta) = \tau\mathbb{E}[l_{1,1}(\theta)] + (1-\tau)\mathbb{E}[l_{2,1}(\theta)]$, and assume the following conditions:

(d3) $\theta_{2-\tau_0}$ is the unique minimizer of $E_{\tau_0}(\theta)$ and lies in the interior of the compact set Θ.

Note that if there is no change, i.e., $\tau = 1, \theta_1$ becomes a unique minimizer of $\mathbb{E}[l_{1,1}(\theta)]$. Especially, $\tilde{\phi}_0$ defined in (8) is the unique minimizer of $\tau_0\mathbb{E}(y_{1,1} - \phi y_{1,0})^2 + (1 - \tau_0)\mathbb{E}(y_{2,1} - \phi y_{2,0})^2$ in AR(1) models.

We can show that under (c1), (c2), (c6) and (d1)–(d3), $\hat{\theta}_n \to \theta_{1-\tau_0}$ a.s., and further, $\sqrt{n}(\hat{\theta}_n - \theta_{1-\tau_0})$ is asymptotically normal by using similar arguments as in the proofs of Theorem 3.1 and 3.2 of Francq and Zakoïan (2004), which implies (C1)(i) and (C2)(i). Concerning (C1)(ii) and (C2)(ii), we refer to the arguments mentioned above Theorem 2. As mentioned therein, they are hard to check analytically in ARMA-GARCH models, too.

3 Simulation study

In this section, we evaluate the performance of the CUSUM test. First, we consider the model:

- ARMA(1,1)-GARCH(1,1) model:

$$\begin{aligned} y_t &= \phi y_{t-1} + \epsilon_t + \psi\epsilon_{t-1}, \\ \epsilon_t &= \sqrt{h_t}\eta_t, \\ h_t &= \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1}. \end{aligned} \tag{15}$$

To investigate empirical size, we consider the following setup:

1. ARMA(1,1)-GARCH(1,1) model with $(\phi, \psi, \omega, \alpha, \beta) = (0.5, 0.1, 0.2, 0.4, 0.2)$.

To examine power, we consider the alternative hypothesis:

$$H_1 : \theta_0 \text{ change to } \theta_1 \text{ occurs at } t = [n/2].$$

For each simulation, we generate sets of $n = 500$ and 800 observations from model (15) with $\eta_t \sim N(0, 1)$. The empirical sizes and powers are calculated at the nominal levels of 0.05 and 0.10 with Gaussian QMLE $\hat{\theta}_n$, which are summarized in Table 1. The figures in the tables denote the proportion of the number of rejections of the null hypothesis, from 500 repetitions.

Table 1 reports the empirical sizes and powers for the ARMA(1,1)-GARCH(1,1) model, showing that the estimate-based CUSUM test $\hat{T}_n^E$ in Lee et al. (2003) has severe size distortions even when the sample size is moderate ($n = 500$) and $\alpha + \beta$ is only 0.6: It is well known that the size distortion becomes more severe as the sum gets closer to 1. The result also reveals that the residual-based CUSUM test $\hat{T}_n^{\mathrm{Res}}$ cannot effectively detect the change in ARMA parameters ϕ and ψ, whereas the $\hat{T}_n^L$ cannot effectively detect the change in GARCH parameters ω, α and β. The score-vector CUSUM test $\hat{T}_n^S$ in Oh and Lee (2017b) also performs as well as $\hat{T}_n$. This finding indicates that the both $\hat{T}_n$ and $\hat{T}_n^S$ are recommendable in dealing with GARCH models with conditional locations and that $\hat{T}_n^{\text{Res}}$ and $\hat{T}_n^L$ perform well for pure GARCH and ARMA models, as might be anticipated.

We compare the performance of $\hat{T}_n$ and $\hat{T}_n^S$ for the threshold AR-GARCH (TAR-GARCH) model and the logistic smooth transition AR-smooth transition GARCH (STAR-STGARCH) model:

- TAR(1)-GARCH(1,1) model:

$$\begin{aligned} y_t &= \big(\phi_1 + \psi_1 y_{t-1}\big) I\big(y_{t-1} > 0\big) + \big(\phi_2 + \psi_2 y_{t-1}\big) I\big(y_{t-1} \leq 0\big) + \epsilon_t, \\ \epsilon_t &= \sqrt{h_t}\eta_t, \\ h_t &= w + a\epsilon_{t-1}^2 + bh_{t-1}. \end{aligned} \tag{16}$$

- Logistic STAR(1)-STGARCH(1,1) model:

$$\begin{aligned} y_t &= \pi_1 + \pi_2 F(y_{t-1}; \gamma_1, c_1) + \big(\varphi_1 + \varphi_2 F(y_{t-1}; \gamma_1, c_1)\big) y_{t-1} + \epsilon_t, \\ \epsilon_t &= \sqrt{h_t}\eta_t, \\ h_t &= \varpi + \big(\alpha_1 + \alpha_2 G(\epsilon_{t-1}; \gamma_2, c_2)\big)\epsilon_{t-1}^2 + \varrho h_{t-1} \end{aligned} \tag{17}$$

with $F(y; \gamma_1, c_1) = [1+\exp(-\gamma_1(y-c_1))]^{-1}$, $G(\epsilon; \gamma_2, c_2) = [1+\exp(-\gamma_2(\epsilon - c_2))]^{-1}$.

An and Huang (1996) and Meitz and Saikkonen (2011) give some sufficient conditions to ensure (M1). We consider the following setup for the null hypothesis:

2. TAR(1)-GARCH(1,1) model with $(\phi_1, \psi_1, \phi_2, \psi_2, w, a, b)$=(0.2, 0.1, 0.1, 0.6, 0.5, 0.1, 0.2).
3. Logistic STAR(1)-STGARCH(1,1) model with $(\pi_1, \pi_2, \varphi_1, \varphi_2, \varpi, \alpha_1, \alpha_2, \varrho)$ =(0.2, 0.4, 0.2, 0.1, 0.2, 0.1, 0.1, 0.2) and transition functions $F(y; 3, 0)$, $G(\epsilon; 6, 0)$.

In these cases, sets of $n = 300$, 500 and 800 observations are generated from models (16) and (17) with $\eta_t \sim N(0, 1)$, $\eta_t \sim \sqrt{4/5}t(10)$, and $\eta_t \sim 0.2N(1.6, 1) + 0.8N(-0.4, 0.2)$. The empirical sizes and powers are calculated at the nominal levels of 0.05 and 0.10 with Gaussian QMLE $\hat{\theta}_n$ and 2,000 repetitions. The results are summarized in Tables 2, 3, 4, 5, 6 and 7.

Tables 2, 3, 4, 5, 6 and 7 report the empirical sizes and powers for the TAR(1)-GARCH(1,1) and Logistic STAR(1)-STGARCH(1,1) models. The $\hat{T}_n^S$ appears to oversize when the sample size is 300 and the error distribution is not normal. Table 2 particularly shows the size-corrected powers. As anticipated, it can be seen that the size-corrected power of $\hat{T}_n$ increases and that of $\hat{T}_n^S$ decreases. Although not reported here, the same pattern can be seen in other cases as well. The results conclude that $\hat{T}_n$ is, on balance, better than $\hat{T}_n^S$.

Tables 8 and 9 report the empirical values of $E_{1,\lambda}$ and $E_{2,\mu}$ in our setup. When there is no parameter change, both $E_{1,\lambda}$ and $E_{2,\mu}$ get closer to 0 as the sample size increases. Moreover, it can be seen that $E_{1,\lambda}$ and $E_{2,\mu}$ have larger values in case the corresponding powers are large.

Table 2 Empirical sizes and powers for the TAR(1)-GARCH(1,1) model with $N(0, 1)$ errors

$(\phi_1, \psi_1, \phi_2, \psi_2, w, a, b)$		$n = 300$		$n = 500$		$n = 800$	
Nominal level		0.05	0.10	0.05	0.10	0.05	0.10
Size	$\hat{T}_n$	**0.027**	**0.068**	**0.033**	**0.076**	**0.040**	**0.081**
(0.2, 0.1, 0.1, 0.6, 0.5, 0.1, 0.2)	$\hat{T}_n^S$	0.106	0.162	0.083	0.137	0.078	0.138
ϕ_1	$\hat{T}_n$	**0.339**	**0.452**	**0.583**	**0.690**	**0.823**	**0.884**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.442	0.588	0.728	0.825	0.944	0.975
Size-corrected	$\hat{T}_n$	**0.414**	**0.526**	**0.640**	**0.731**	**0.841**	**0.894**
Power	$\hat{T}_n^S$	0.198	0.395	0.574	0.724	0.696	0.758
ψ_1	$\hat{T}_n$	**0.244**	**0.358**	**0.456**	**0.579**	**0.727**	**0.813**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.282	0.421	0.552	0.685	0.849	0.915
Size-corrected	$\hat{T}_n$	**0.327**	**0.428**	**0.524**	**0.618**	**0.754**	**0.828**
Power	$\hat{T}_n^S$	0.100	0.249	0.404	0.591	0.708	0.839
ϕ_2	$\hat{T}_n$	**0.230**	**0.373**	**0.517**	**0.655**	**0.836**	**0.910**
$0.1 \longrightarrow -0.3$	$\hat{T}_n^S$	0.421	0.573	0.746	0.835	0.953	0.976
Size-corrected	$\hat{T}_n$	**0.331**	**0.471**	**0.590**	**0.702**	**0.863**	**0.920**
Power	$\hat{T}_n^S$	0.182	0.387	0.593	0.756	0.724	0.770
ψ_2	$\hat{T}_n$	**0.162**	**0.243**	**0.308**	**0.428**	**0.493**	**0.605**
$0.6 \longrightarrow 0.2$	$\hat{T}_n^S$	0.230	0.349	0.390	0.539	0.641	0.761
Size-corrected	$\hat{T}_n$	**0.217**	**0.398**	**0.369**	**0.468**	**0.531**	**0.632**
Power	$\hat{T}_n^S$	0.077	0.208	0.258	0.443	0.480	0.673

Table 2 continued

$(\phi_1, \psi_1, \phi_2, \psi_2, w, a, b)$		$n = 300$		$n = 500$		$n = 800$	
Nominal level		0.05	0.10	0.05	0.10	0.05	0.10
w	$\hat{T}_n$	**0.335**	**0.465**	**0.619**	**0.737**	**0.780**	**0.853**
$0.5 \longrightarrow 0.9$	$\hat{T}_n^S$	0.215	0.318	0.409	0.547	0.619	0.734
Size-corrected	$\hat{T}_n$	**0.419**	**0.560**	**0.681**	**0.781**	**0.805**	**0.868**
Power	$\hat{T}_n^S$	0.066	0.189	0.282	0.457	0.487	0.647
a	$\hat{T}_n$	**0.202**	**0.321**	**0.453**	**0.590**	**0.752**	**0.847**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.117	0.228	0.278	0.432	0.600	0.757
Size-corrected	$\hat{T}_n$	**0.287**	**0.404**	**0.526**	**0.650**	**0.784**	**0.807**
Power	$\hat{T}_n^S$	0.025	0.102	0.163	0.340	0.435	0.675
b	$\hat{T}_n$	**0.308**	**0.427**	**0.447**	**0.557**	**0.549**	**0.654**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.216	0.322	0.280	0.403	0.384	0.480
Size-corrected	$\hat{T}_n$	**0.385**	**0.507**	**0.509**	**0.610**	**0.576**	**0.683**
Power	$\hat{T}_n^S$	0.079	0.194	0.194	0.315	0.263	0.406

Table 3 Empirical sizes and powers for the TAR(1)-GARCH(1,1) model with $\sqrt{4/5}t(10)$ errors

$(\phi_1, \psi_1, \phi_2, \psi_2, w, a, b)$		$n = 300$		$n = 500$		$n = 800$	
Nominal level		0.05	0.10	0.05	0.10	0.05	0.10
Size	$\hat{T}_n$	**0.024**	**0.061**	**0.035**	**0.072**	**0.040**	**0.089**
(0.2,0.1,0.1,0.6,0.5,0.1,0.2)	$\hat{T}_n^S$	0.117	0.167	0.098	0.153	0.089	0.149
ϕ_1	$\hat{T}_n$	**0.383**	**0.492**	**0.604**	**0.695**	**0.844**	**0.894**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.491	0.627	0.745	0.853	0.948	0.975
ψ_1	$\hat{T}_n$	**0.249**	**0.363**	**0.476**	**0.583**	**0.704**	**0.794**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.327	0.458	0.565	0.694	0.847	0.909
ϕ_2	$\hat{T}_n$	**0.233**	**0.364**	**0.501**	**0.646**	**0.823**	**0.895**
$0.1 \longrightarrow -0.3$	$\hat{T}_n^S$	0.446	0.596	0.732	0.842	0.964	0.981
ψ_2	$\hat{T}_n$	**0.168**	**0.252**	**0.273**	**0.381**	**0.470**	**0.586**
$0.6 \longrightarrow 0.2$	$\hat{T}_n^S$	0.256	0.367	0.389	0.509	0.645	0.761
w	$\hat{T}_n$	**0.230**	**0.363**	**0.479**	**0.620**	**0.731**	**0.815**
$0.5 \longrightarrow 0.9$	$\hat{T}_n^S$	0.181	0.298	0.323	0.448	0.539	0.659
a	$\hat{T}_n$	**0.117**	**0.213**	**0.274**	**0.415**	**0.492**	**0.615**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.109	0.193	0.185	0.307	0.370	0.536
b	$\hat{T}_n$	**0.250**	**0.372**	**0.421**	**0.536**	**0.545**	**0.647**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.188	0.286	0.303	0.414	0.404	0.504

Overall, our findings confirm the validity of the proposed residual-based CUSUM test and its advantages over the score vector-based CUSUM test in terms of simplicity and performance.

4 Real data analysis

In this section, we apply our test to daily Dow30 data. We analyze the log returns of Dow30 data (Fig. 1), from December 15, 2014 to March 17, 2017 with 568 observations. Inspection of SACF, SPACF, AIC and BIC results suggests that an AR(1) model is suitable to the data. Moreover, Fig. 1 show that the returns have some volatility clustering phenomenon. Since the Ljung–Box and LM-ARCH tests based on the AR(1) residuals reveal that the GARCH(1,1) model is reasonable for this series, we fit an AR(1)-GARCH(1,1) model to the data and obtain the estimated model as follows:

$$\begin{aligned} y_t &= -0.05901(0.04693)y_{t-1} + \epsilon_t, \\ h_t &= 7.816 \times 10^{-6}(2.651 \times 10^{-6}) \\ &\quad + 0.2110(0.05182)\epsilon_{t-1}^2 + 0.6786(0.07452)h_{t-1}. \end{aligned} \tag{18}$$

The figures in parentheses denote the corresponding standard errors of the parameter estimates.

Table 4 Empirical sizes and powers for the TAR(1)-GARCH(1,1) model with $0.2N(1.6, 1) + 0.8N(-0.4, 0.2)$ errors

$(\phi_1, \psi_1, \phi_2, \psi_2, w, a, b)$		$n = 300$		$n = 500$		$n = 800$	
Nominal level		0.05	0.10	0.05	0.10	0.05	0.10
Size	$\hat{T}_n$	**0.034**	**0.073**	**0.041**	**0.092**	**0.047**	**0.084**
(0.2, 0.1, 0.1, 0.6, 0.5, 0.1, 0.2)	$\hat{T}_n^S$	0.188	0.254	0.132	0.194	0.111	0.165
ϕ_1	$\hat{T}_n$	**0.408**	**0.523**	**0.651**	**0.752**	**0.864**	**0.918**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.618	0.727	0.845	0.917	0.984	0.994
ψ_1	$\hat{T}_n$	**0.324**	**0.441**	**0.598**	**0.710**	**0.833**	**0.895**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.550	0.685	0.845	0.917	0.986	0.995
ϕ_2	$\hat{T}_n$	**0.191**	**0.297**	**0.408**	**0.546**	**0.680**	**0.807**
$0.1 \longrightarrow -0.3$	$\hat{T}_n^S$	0.664	0.771	0.903	0.951	0.995	0.998
ψ_2	$\hat{T}_n$	**0.086**	**0.144**	**0.120**	**0.209**	**0.220**	**0.312**
$0.6 \longrightarrow 0.2$	$\hat{T}_n^S$	0.284	0.398	0.351	0.490	0.504	0.632
w	$\hat{T}_n$	**0.182**	**0.291**	**0.387**	**0.530**	**0.686**	**0.779**
$0.5 \longrightarrow 0.9$	$\hat{T}_n^S$	0.374	0.497	0.574	0.691	0.820	0.873
a	$\hat{T}_n$	**0.080**	**0.145**	**0.139**	**0.236**	**0.295**	**0.417**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.166	0.251	0.257	0.381	0.466	0.622
b	$\hat{T}_n$	**0.237**	**0.350**	**0.435**	**0.566**	**0.618**	**0.710**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.419	0.541	0.605	0.684	0.718	0.773

Table 5 Empirical sizes and powers for the Logistic STAR(1)-STGARCH(1,1) model with $N(0, 1)$ errors

$(\pi_1, \pi_2, \varphi_1, \varphi_2, \varpi, \alpha_1, \alpha_2, \varrho)$		$n = 300$		$n = 500$		$n = 800$	
Nominal level		0.05	0.10	0.05	0.10	0.05	0.10
Size	$\hat{T}_n$	**0.031**	**0.071**	**0.040**	**0.079**	**0.054**	**0.095**
(0.2,0.4,0.2,0.1,0.2,0.1,0.1,0.2)	$\hat{T}_n^S$	0.155	0.217	0.099	0.152	0.065	0.118
π_1	$\hat{T}_n$	**0.317**	**0.433**	**0.556**	**0.669**	**0.815**	**0.871**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.471	0.596	0.733	0.826	0.933	0.972
π_2	$\hat{T}_n$	**0.062**	**0.115**	**0.134**	**0.219**	**0.197**	**0.292**
$0.4 \longrightarrow 0.8$	$\hat{T}_n^S$	0.198	0.284	0.240	0.355	0.348	0.478
φ_1	$\hat{T}_n$	**0.859**	**0.910**	**0.994**	**0.998**	**1.000**	**1.000**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.703	0.813	0.969	0.986	1.000	1.000
φ_2	$\hat{T}_n$	**0.490**	**0.613**	**0.828**	**0.894**	**0.972**	**0.986**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.393	0.529	0.674	0.796	0.926	0.961
ϖ	$\hat{T}_n$	**0.446**	**0.505**	**0.396**	**0.465**	**0.399**	**0.484**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.621	0.717	0.725	0.823	0.812	0.883
α_1	$\hat{T}_n$	**0.298**	**0.442**	**0.630**	**0.756**	**0.888**	**0.943**
$0.1 \longrightarrow 0.6$	$\hat{T}_n^S$	0.143	0.261	0.336	0.504	0.726	0.861
α_2	$\hat{T}_n$	**0.081**	**0.140**	**0.156**	**0.240**	**0.285**	**0.390**
$0.1 \longrightarrow 0.6$	$\hat{T}_n^S$	0.103	0.184	0.142	0.252	0.252	0.391
ϱ	$\hat{T}_n$	**0.185**	**0.261**	**0.167**	**0.256**	**0.220**	**0.318**
$0.2 \longrightarrow 0.7$	$\hat{T}_n^S$	0.362	0.473	0.371	0.504	0.387	0.539

Table 6 Empirical sizes and powers for the Logistic STAR(1)-STGARCH(1,1) model with $\sqrt{4/5}t(10)$ errors

$(\pi_1, \pi_2, \varphi_1, \varphi_2, \varpi, \alpha_1, \alpha_2, \varrho)$		$n=300$		$n=500$		$n=800$	
Nominal level		0.05	0.10	0.05	0.10	0.05	0.10
Size	$\hat{T}_n$	**0.035**	**0.072**	**0.037**	**0.076**	**0.040**	**0.083**
(0.2,0.4,0.2,0.1,0.2,0.1,0.1,0.2)	$\hat{T}_n^S$	0.180	0.251	0.112	0.174	0.077	0.132
π_1	$\hat{T}_n$	**0.353**	**0.469**	**0.507**	**0.626**	**0.815**	**0.881**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.524	0.647	0.693	0.798	0.927	0.963
π_2	$\hat{T}_n$	**0.086**	**0.152**	**0.136**	**0.225**	**0.221**	**0.319**
$0.4 \longrightarrow 0.8$	$\hat{T}_n^S$	0.240	0.333	0.257	0.374	0.388	0.519
φ_1	$\hat{T}_n$	**0.870**	**0.922**	**0.978**	**0.991**	**1.000**	**1.000**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.764	0.853	0.931	0.955	1.000	1.000
φ_2	$\hat{T}_n$	**0.454**	**0.589**	**0.788**	**0.868**	**0.960**	**0.978**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.438	0.578	0.667	0.788	0.907	0.949
ϖ	$\hat{T}_n$	**0.479**	**0.566**	**0.534**	**0.599**	**0.504**	**0.596**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.577	0.682	0.742	0.822	0.830	0.886
α_1	$\hat{T}_n$	**0.170**	**0.306**	**0.391**	**0.533**	**0.660**	**0.772**
$0.1 \longrightarrow 0.6$	$\hat{T}_n^S$	0.148	0.247	0.230	0.384	0.466	0.636
α_2	$\hat{T}_n$	**0.056**	**0.102**	**0.100**	**0.180**	**0.179**	**0.282**
$0.1 \longrightarrow 0.6$	$\hat{T}_n^S$	0.116	0.186	0.114	0.193	0.161	0.279
ϱ	$\hat{T}_n$	**0.206**	**0.289**	**0.187**	**0.283**	**0.267**	**0.381**
$0.2 \longrightarrow 0.7$	$\hat{T}_n^S$	0.378	0.486	0.380	0.496	0.408	0.549

Table 7 Empirical sizes and powers for the logistic STAR(1)-STGARCH(1,1) model with $0.2N(1.6, 1) + 0.8N(-0.4, 0.2)$ errors

$(\pi_1, \pi_2, \varphi_1, \varphi_2, \varpi, \alpha_1, \alpha_2, \varrho)$		$n = 300$		$n = 500$		$n = 800$	
Nominal level		0.05	0.10	0.05	0.10	0.05	0.10
Size	$\hat{T}_n$	**0.038**	**0.080**	**0.044**	**0.083**	**0.045**	**0.097**
(0.2,0.4,0.2,0.1,0.2,0.1,0.1,0.2)	$\hat{T}_n^S$	0.315	0.392	0.221	0.293	0.171	0.232
π_1	$\hat{T}_n$	**0.706**	**0.787**	**0.915**	**0.950**	**0.994**	**0.996**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.842	0.900	0.966	0.987	1.000	1.000
π_2	$\hat{T}_n$	**0.054**	**0.106**	**0.062**	**0.102**	**0.065**	**0.111**
$0.4 \longrightarrow 0.8$	$\hat{T}_n^S$	0.359	0.440	0.287	0.378	0.272	0.362
φ_1	$\hat{T}_n$	**0.955**	**0.975**	**0.999**	**1.000**	**1.000**	**1.000**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.961	0.987	1.000	1.000	1.000	1.000
φ_2	$\hat{T}_n$	**0.616**	**0.741**	**0.895**	**0.944**	**0.989**	**0.995**
$0.1 \longrightarrow 0.5$	$\hat{T}_n^S$	0.745	0.830	0.932	0.967	0.995	0.998
ϖ	$\hat{T}_n$	**0.671**	**0.739**	**0.773**	**0.805**	**0.794**	**0.836**
$0.2 \longrightarrow 0.6$	$\hat{T}_n^S$	0.850	0.905	0.955	0.972	0.971	0.980
α_1	$\hat{T}_n$	**0.146**	**0.239**	**0.310**	**0.428**	**0.548**	**0.672**
$0.1 \longrightarrow 0.6$	$\hat{T}_n^S$	0.263	0.358	0.316	0.441	0.528	0.665
α_2	$\hat{T}_n$	**0.062**	**0.116**	**0.124**	**0.189**	**0.200**	**0.300**
$0.1 \longrightarrow 0.6$	$\hat{T}_n^S$	0.223	0.309	0.208	0.310	0.293	0.430
ϱ	$\hat{T}_n$	**0.363**	**0.458**	**0.465**	**0.573**	**0.595**	**0.689**
$0.2 \longrightarrow 0.7$	$\hat{T}_n^S$	0.715	0.787	0.776	0.840	0.870	0.918

Table 8 Empirical value of $E_{1,\lambda}$ and $E_{2,\mu}$ for the TAR(1)-GARCH(1,1) model

Parameter change	$n = 300$		$n = 500$		$n = 800$	
	$\bar{\hat{E}}_{1,\lambda}$	$\bar{\hat{E}}_{2,\mu}$	$\bar{\hat{E}}_{1,\lambda}$	$\bar{\hat{E}}_{2,\mu}$	$\bar{\hat{E}}_{1,\lambda}$	$\bar{\hat{E}}_{2,\mu}$
No change	0.0739	0.3149	0.0584	0.2524	0.0468	0.2048
ϕ_1	0.1506	0.2245	0.1414	0.1582	0.1349	0.1163
ψ_1	0.1485	0.2417	0.1377	0.1690	0.1306	0.1213
ϕ_2	0.1892	0.2400	0.1780	0.1686	0.1696	0.1218
ψ_2	0.0738	0.2711	0.0646	0.2045	0.0592	0.1468
w	0.0643	0.4811	0.0436	0.4559	0.0315	0.4343
a	0.0826	0.4275	0.0606	0.3917	0.0412	0.3642
b	0.0697	0.4904	0.0503	0.4413	0.0371	0.3960

Table 9 Empirical value of $E_{1,\lambda}$ and $E_{2,\mu}$ for the Logistic STAR(1)-STGARCH(1,1) model

Parameter change	$n = 300$		$n = 500$		$n = 800$	
	$\bar{\hat{E}}_{1,\lambda}$	$\bar{\hat{E}}_{2,\mu}$	$\bar{\hat{E}}_{1,\lambda}$	$\bar{\hat{E}}_{2,\mu}$	$\bar{\hat{E}}_{1,\lambda}$	$\bar{\hat{E}}_{2,\mu}$
No change	0.0449	0.3318	0.0360	0.2534	0.0289	0.2019
π_1	0.0921	0.2478	0.0840	0.1617	0.0825	0.1111
π_2	0.0572	0.3207	0.0481	0.2361	0.0411	0.1768
φ_1	0.1768	0.2309	0.1732	0.1419	0.1724	0.1065
φ_2	0.1241	0.2316	0.1196	0.1492	0.1167	0.1104
ϖ	0.0410	1.4378	0.0355	1.5453	0.0323	1.4453
α_1	0.0442	0.4555	0.0302	0.4143	0.0211	0.3876
α_2	0.0450	0.3567	0.0333	0.3021	0.0245	0.2680
ϱ	0.0446	0.9659	0.0342	0.7200	0.0266	0.5994

Next, we perform a parameter change test for $(\phi, \omega, \alpha, \beta)$ and get p value 0.046, which rejects the null hypothesis of no parameter changes at the nominal level of 0.05. This result coincides with that of the score-based CUSUM test.

Finally, we perform $\hat{T}_n^L$ and $\hat{T}_n^{\text{Res}}$ to check whether the change occurs in the location or scale parameters. The location parameter change test ($\hat{T}_n^L$) has p value 0.4354, whereas the residual-based CUSUM test ($\hat{T}_n^{\text{Res}}$) has p value 0.016, which rejects the null of no parameter changes at the nominal level of 0.05. From this, we can conclude that only a scale parameter change occurs. Using $\hat{T}_n^{\text{Res}}$, we can also see that the change point is located at March 1, 2016 (dashed lines in Fig. 1). A visual inspection of Fig. 1 clearly shows a dispersion change on this date.

The first subseries in the pre-change period, from December 15, 2014 to March 1, 2016, follows the AR(1)-GARCH(1,1) model with

$$\begin{aligned} h_t &= 1.338 \times 10^{-5}(8.078 \times 10^{-6}) \\ &\quad + 0.1887(0.07485)\epsilon_{t-1}^2 + 0.6817(0.1333)h_{t-1}, \end{aligned}$$

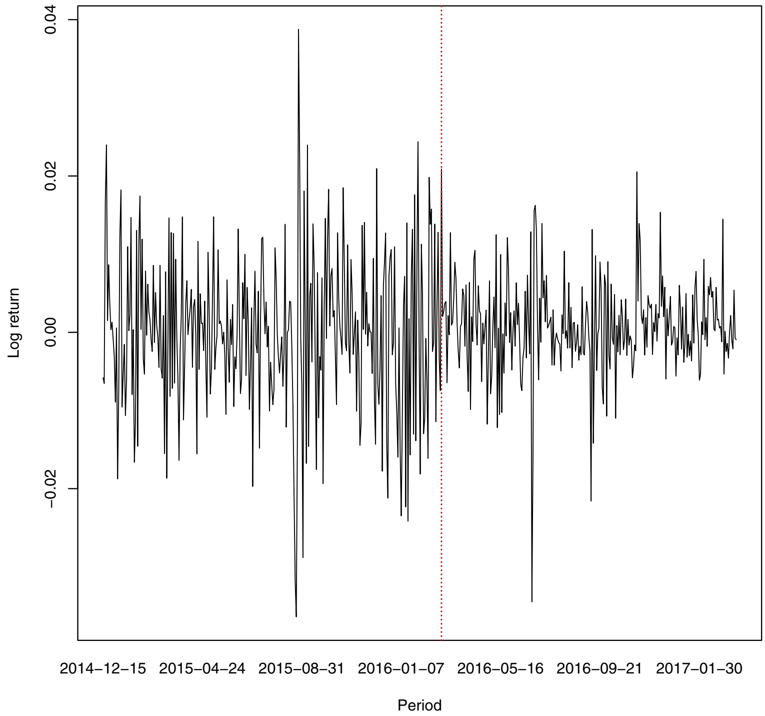

Fig. 1 Log-return of Dow30 data from Dec 15, 2014 to Mar 17, 2017

and the second subseries in the post-change period follows the AR(1)-GARCH(1,1) model with

$$h_t = 2.088 \times 10^{-5}(5.572 \times 10^{-6}) + 0.2637(0.09123)\epsilon_{t-1}^2 + 0.1410(0.1862)h_{t-1}.$$

This result also confirms that the parameter experiences a significant change particularly in β.

5 Concluding remarks

We proposed a modified residual-based CUSUM test to detect a parameter change in location-scale heteroscedastic time series models. We derived their limiting null distributions as the sup of the sum of squares of independent Brownian bridges. We also demonstrated the validity of the proposed CUSUM tests though Monte Carlo simulations and performed a data analysis using Dow30 dataset. Our findings confirmed the validity of the newly proposed residual-based CUSUM test as an "omnibus" test to detect a parameter change in location-scale GARCH-type models. In this study, we only focused on the retrospective change point detection problem and leave the "on-line monitoring problem" as a future project.

6 Proofs

For notational convenience, we express $f_1(\theta)f_2(\theta)$ as $f_1f_2(\theta)$ for any functions f_1, f_2: Further, if f_1 is only a function of μ, $f_1(\theta)$ indicates $f_1(\mu)$. We use $\partial_\theta f = \frac{\partial f(\theta)}{\partial\theta}$ to stand for the derivatives of functions f. Let $\epsilon_t(\mu) = y_t - g_t(\mu)$, $\eta_t(\theta) = \frac{\epsilon_t}{\sqrt{h_t}}(\theta)$, $\tilde{\epsilon}_t(\mu) = y_t - \tilde{g}_t(\mu)$: Notice that $\eta_t = \eta_t(\theta_0)$.

Lemma 1 *Under the same conditions in Theorem 1, we have*

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\left|\sum_{t=1}^{k}\left((\tilde{g}_t\tilde{\eta}_t)\left(\hat{\theta}_n\right) - (g_t\eta_t)(\theta_0)\right) - \frac{k}{n}\sum_{t=1}^{n}\left((\tilde{g}_t\tilde{\eta}_t)\left(\hat{\theta}_n\right) - (g_t\eta_t)(\theta_0)\right)\right|$$
$$= o_P(1).$$

Proof We can express $(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n)$ as follows:

$$\begin{aligned}(\tilde{g}_t\tilde{\eta}_t)\left(\hat{\theta}_n\right) &= (g_t\eta_t)(\theta_0) + \left\{(\tilde{g}_t\tilde{\eta}_t)\left(\hat{\theta}_n\right) - (g_t\eta_t)\left(\hat{\theta}_n\right)\right\} \\ &\quad + \left\{(g_t\eta_t)\left(\hat{\theta}_n\right) - (g_t\eta_t)(\theta_0)\right\} \\ &:= (g_t\eta_t)(\theta_0) + I_{1,t} + I_{2,t}.\end{aligned}$$

We first show that

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\left|\sum_{t=1}^{k} I_{1,t} - \frac{k}{n}\sum_{t=1}^{n} I_{1,t}\right| = o_P(1).$$

Note that

$$\begin{aligned}I_{1,t} &= \frac{g_t(\tilde{\epsilon}_t - \epsilon_t)}{\sqrt{\tilde{h}_t}}(\hat{\theta}_n) + \frac{(\tilde{g}_t - g_t)\epsilon_t}{\sqrt{\tilde{h}_t}}\left(\hat{\theta}_n\right) \\ &\quad + \frac{(\tilde{g}_t - g_t)(\tilde{\epsilon}_t - \epsilon_t)}{\sqrt{\tilde{h}_t}}\left(\hat{\theta}_n\right) + g_t\epsilon_t\left(\frac{1}{\sqrt{\tilde{h}_t}} - \frac{1}{\sqrt{h_t}}\right)\left(\hat{\theta}_n\right) \\ &:= J_{1,t} + J_{2,t} + J_{3,t} + J_{4,t}.\end{aligned}$$

We first deal with $J_{1,t}$. For any $\delta > 0$ and neighborhood $N(\theta_0) \subset \Theta$,

$$\mathrm{P}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left|J_{1,t}\right| > \delta\right) \le \mathrm{P}\left(\hat{\theta}_n \in \Theta \setminus N(\theta_0)\right) + P\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\sup_{\theta\in\Theta}\left|g_t\left(\frac{\tilde{\epsilon}_t - \epsilon_t}{\sqrt{\tilde{h}_t}}\right)(\theta)\right| > \delta\right).$$

Note that

$$\sum_{t=1}^{n}\sup_{\theta\in\Theta}\left|g_t\left(\frac{\tilde{\epsilon}_t - \epsilon_t}{\sqrt{\tilde{h}_t}}\right)(\theta)\right| \le \frac{1}{\sqrt{\underline{h}}}\sum_{t=1}^{n}\sup_{\theta\in\Theta}\left|g_t(\tilde{g}_t - g_t)(\mu)\right|.$$

Since $\sup_{\theta\in\Theta}|\tilde{g}_t(\mu) - g_t(\mu)| \le V_g(y_{t-1},\ldots)\cdot\rho^{\mathrm{T}}$ a.s. and $\mathbb{E}\sup_{\theta\in\Theta}|g_t(\mu)|^4 < \infty$ implies $\mathbb{E}\log^+\sup_{\theta\in\Theta}|g_t(\mu)| < \infty$, Lemma 2.1 of Straumann and Mikosch (2006) implies $\sum_{t=1}^{\infty}\sup_{\theta\in\Theta}|g_t(\tilde{g}_t - g_t)(\mu)|$ converges a.s. Thus, using the fact $\hat{\theta}_n = \theta_0 + O_{\mathrm{P}}(1/\sqrt{n})$, we get

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\sum_{t=1}^{k}|J_{1,t}| \le \frac{1}{\sqrt{n}}\sum_{t=1}^{n}|J_{1,t}| = o_{\mathrm{P}}(1). \tag{19}$$

Next, we handle $J_{i,t}$, $i = 2, 3, 4$. Notice that

$$\left|\frac{(\tilde{g}_t - g_t)\epsilon_t}{\sqrt{\tilde{h}_t}}(\theta)\right| \le \frac{1}{\sqrt{\underline{h}}}\left|(\tilde{g}_t - g_t)\epsilon_t(\theta)\right|,$$

$$\left|\frac{(\tilde{g}_t - g_t)(\tilde{\epsilon}_t - \epsilon_t)}{\sqrt{\tilde{h}_t}}(\theta)\right| \le \frac{1}{\sqrt{\underline{h}}}\left|(\tilde{g}_t - g_t)(\theta)\right|^2$$

and

$$\left|g_t\epsilon_t\left(\frac{1}{\sqrt{\tilde{h}_t}} - \frac{1}{\sqrt{h_t}}\right)(\theta)\right| \le \frac{1}{2\underline{h}^{3/2}}\left|y_t g_t(\tilde{h}_t - h_t)(\theta) + g_t^2(\tilde{h}_t - h_t)(\theta)\right|.$$

Thus, similar to (19), we can easily see that

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\sum_{t=1}^{k}|J_{i,t}| \le \frac{1}{\sqrt{n}}\sum_{t=1}^{n}|J_{i,t}| = o_{\mathrm{P}}(1), \quad i = 2, 3, 4. \tag{20}$$

Finally, we deal with $I_{2,t}$. By the mean value theorem, we have

$$\begin{aligned}(g_t\eta_t)(\hat{\theta}_n) - (g_t\eta_t)(\theta_0) &= (\hat{\theta}_n - \theta_0)^{\mathrm{T}} \left(\frac{\partial_\theta g_t \epsilon_t}{\sqrt{h_t}} + \frac{g_t \partial_\theta \epsilon_t}{\sqrt{h_t}} + \frac{g_t \epsilon_t \partial_\theta h_t}{2h_t^{3/2}} \right)(\bar{\theta}_n) \\ &:= (\hat{\theta}_n - \theta_0)^{\mathrm{T}} D_t(\bar{\theta}_n),\end{aligned}$$

where $\bar{\theta}_n$ is an intermediate point between $\hat{\theta}_n$ and θ_0. For $\zeta > 0$, let $N(\theta_0, \zeta) = \{\theta : ||\theta - \theta_0|| \le \zeta\}$. Note that for any $\delta > 0$,

$$\begin{aligned}&\mathrm{P}\left(\max_{1\le k\le n} \left\| \frac{1}{k}\sum_{t=1}^{k} D_t(\bar{\theta}_n) - \frac{1}{k}\sum_{t=1}^{k} D_t(\theta_0) \right\| \ge \delta \right) \\ &\le \mathrm{P}\left(\bar{\theta}_n \in \Theta \setminus N(\theta_0, K/\sqrt{n})\right) \\ &\quad + \mathrm{P}\left(\max_{1\le k\le n} \frac{1}{k}\sum_{t=1}^{k} \sup_{\theta\in N(\theta_0, K/\sqrt{n})} \|D_t(\theta) - D_t(\theta_0)\| \ge \delta \right).\end{aligned} \tag{21}$$

Then, using the fact $\partial_\theta \epsilon_t(\mu) = -\partial_\theta g_t(\mu)$, (A4), (A8) and Hölder's inequality, one can readily check that $\mathbb{E}\sup_{\theta\in N(\theta_0)} \left\| D_t(\theta) \right\| < \infty$ and $D_t(\theta)$ is stationary and ergodic, which implies that as $k \to \infty$,

$$\frac{1}{k}\sum_{t=1}^{k} \sup_{\theta\in N(\theta_0,\zeta)} \left\| D_t(\theta) - D_t(\theta_0) \right\| \longrightarrow \mathbb{E} \sup_{\theta\in N(\theta_0,\zeta)} \left\| D_t(\theta) - D_t(\theta_0) \right\|, \quad a.s.$$

Hence, for any $\delta > 0$, by choosing sufficiently large $K > 0$, we can get

$$\lim_{n\to\infty} \max_{1\le k\le n} \frac{1}{k}\sum_{t=1}^{k} \sup_{\theta\in N(\theta_0, K/\sqrt{n})} \left\| D_t(\theta) - D_t(\theta_0) \right\| < \delta \;\; a.s. \tag{22}$$

Then, combining (21), (22) and the fact that $\hat{\theta}_n = \theta_0 + O_{\mathrm{P}}(1/\sqrt{n})$, we have

$$\begin{aligned}\left\| \frac{1}{n}\sum_{t=1}^{n} D_t(\bar{\theta}_n) - \frac{1}{n}\sum_{t=1}^{n} D_t(\theta_0) \right\| &\le \max_{1\le k\le n} \left\| \frac{1}{k}\sum_{t=1}^{k} D_t(\bar{\theta}_n) - \frac{1}{k}\sum_{t=1}^{k} D_t(\theta_0) \right\| \\ &= o_{\mathrm{P}}(1),\end{aligned} \tag{23}$$

which indicates

$$\frac{1}{\sqrt{n}} \max_{1\le k\le n} \left| \sum_{t=1}^{k} I_{2,t} - \frac{k}{n}\sum_{t=1}^{n} I_{2,t} \right| = o_{\mathrm{P}}(1) \tag{24}$$

with the fact $\hat{\theta}_n = \theta + O_{\mathrm{P}}(1/\sqrt{n})$. This together with (19) and (20) validates the lemma. □

Lemma 2 *Under the same conditions in Theorem 1, we have*

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\left|\sum_{t=1}^{k}\left(\tilde{\eta}_t^2(\hat{\theta}_n)-\eta_t^2\right)-\frac{k}{n}\sum_{t=1}^{n}\left(\tilde{\eta}_t^2(\hat{\theta}_n)-\eta_t^2\right)\right|=o_{\mathrm{P}}(1).$$

Proof We express

$$\begin{aligned}\tilde{\eta}_t^2(\hat{\theta}_n)&=\eta_t^2+\left(\frac{\tilde{\epsilon}_t^2}{\tilde{h}_t}-\frac{\epsilon_t^2}{h_t}\right)(\hat{\theta}_n)+\left(\frac{\epsilon_t^2}{h_t}(\hat{\theta}_n)-\frac{\epsilon_t^2}{h_t}(\theta_0)\right)\\&=\eta_t^2+R_{1,t}+R_{2,t}.\end{aligned}$$

Note that

$$R_{1,t}=\frac{2\epsilon_t(\tilde{g}_t-g_t)}{\tilde{h}_t}(\hat{\theta}_n)+\frac{(\tilde{g}_t-g_t)^2}{\tilde{h}_t}(\hat{\theta}_n)+\epsilon_t^2\left(\frac{1}{\tilde{h}_t}-\frac{1}{h_t}\right)(\hat{\theta}_n)$$

and

$$\left|\frac{2\epsilon_t(\tilde{g}_t-g_t)}{\tilde{h}_t}(\theta)\right|\le\frac{2}{\underline{h}}\left|\epsilon_t(\tilde{g}_t-g_t)(\theta)\right|,$$

$$\left|\frac{(\tilde{g}_t-g_t)^2}{\tilde{h}_t}(\theta)\right|\le\frac{1}{\underline{h}}|(\tilde{g}_t-g_t)(\theta)|^2$$

and

$$\left|\epsilon_t^2\left(\frac{1}{\tilde{h}_t}-\frac{1}{h_t}\right)(\theta)\right|\le\frac{1}{\underline{h}^2}\left|\epsilon_t^2(\tilde{h}_t-h_t)(\theta)\right|.$$

Thus, similar to (19), one can easily show that

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\sum_{t=1}^{k}|R_{1,t}|\le\frac{1}{\sqrt{n}}\sum_{t=1}^{n}|R_{1,t}|=o_{\mathrm{P}}(1). \tag{25}$$

Meanwhile, by the mean value theorem, we have

$$\frac{\epsilon_t^2}{h_t}(\hat{\theta}_n)-\frac{\epsilon_t^2}{h_t}(\theta_0)=(\hat{\theta}_n-\theta_0)^{\mathrm{T}}\left(\frac{2\epsilon_t\partial_\theta\epsilon_t}{h_t}-\frac{\epsilon_t^2\partial_\theta h_t}{h_t^2}\right)(\bar{\theta}_n).$$

Then, since $\mathbb{E}\sup_{\theta\in N(\theta_0)}\left\|\left(\frac{2\epsilon_t\partial_\theta\epsilon_t}{h_t}-\frac{\epsilon_t^2\partial_\theta h_t}{h_t^2}\right)(\theta)\right\|<\infty$ and $\left(\frac{2\epsilon_t\partial_\theta\epsilon_t}{h_t}-\frac{\epsilon_t^2\partial_\theta h_t}{h_t^2}\right)(\theta)$ is stationary and ergodic, we have similar to (24),

$$\frac{1}{\sqrt{n}}\max_{1\le k\le n}\left|\sum_{t=1}^{k}R_{2,t}-\frac{k}{n}\sum_{t=1}^{n}R_{2,t}\right|=o_{\mathrm{P}}(1).$$

This together with (25) asserts the lemma. □

Proof of Theorem 1 We can write

$$\begin{aligned}
&\Sigma_0^{-\frac{1}{2}}\frac{1}{\sqrt{n}}\left(\sum_{t=1}^{k}\tilde{U}_t(\hat{\theta}_n)-\frac{k}{n}\sum_{t=1}^{n}\tilde{U}_t(\hat{\theta}_n)\right)\\
&=\Sigma_0^{-\frac{1}{2}}\frac{1}{\sqrt{n}}\left\{\sum_{t=1}^{k}\left(\tilde{U}_t(\hat{\theta}_n)-\mathbb{E}U_t(\theta_0)\right)-\frac{k}{n}\sum_{t=1}^{n}\left(\tilde{U}_t(\hat{\theta}_n)-\mathbb{E}U_t(\theta_0)\right)\right\}\\
&=\Sigma_0^{-\frac{1}{2}}\frac{1}{\sqrt{n}}\left\{\sum_{t=1}^{k}\left(U_t(\theta_0)-\mathbb{E}U_t(\theta_0)\right)-\frac{k}{n}\sum_{t=1}^{n}\left(U_t(\theta_0)-\mathbb{E}U_t(\theta_0)\right)\right\}\\
&\quad+\Sigma_0^{-\frac{1}{2}}\frac{1}{\sqrt{n}}\left\{\sum_{t=1}^{k}\left(\tilde{U}_t(\hat{\theta}_n)-U_t(\theta_0)\right)-\frac{k}{n}\sum_{t=1}^{n}\left(\tilde{U}_t(\hat{\theta}_n)-U_t(\theta_0)\right)\right\},
\end{aligned}$$

wherein the second term of the last equality is asymptotically negligible due to Lemmas 1 and 2. Then, the theorem is obtained by using the functional central limit theorem for martingales (Theorem 23.1 of Billingsley (1968)) and the Cramér-Wold device. □

Lemma 3 *Under the same conditions in Theorem 1, we have that as $n\to\infty$,*

$$\hat{\Sigma}_n\xrightarrow{\mathrm{P}}\Sigma_0.$$

Proof We first show that

$$\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n)\xrightarrow{\mathrm{P}}\mathbb{E}\left(g_t^2\eta_t^2\right)(\theta_0). \tag{26}$$

Note that

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n)&=\left(\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n)-\frac{1}{n}\sum_{t=1}^{n}\left(g_t^2\eta_t^2\right)(\hat{\theta}_n)\right)\\
&\quad+\left(\frac{1}{n}\sum_{t=1}^{n}\left(g_t^2\eta_t^2\right)(\hat{\theta}_n)-\frac{1}{n}\sum_{t=1}^{n}\left(g_t^2\eta_t^2\right)(\theta_0)\right)\\
&\quad+\frac{1}{n}\sum_{t=1}^{n}\left(g_t^2\eta_t^2\right)(\theta_0)
\end{aligned}$$

and

$$\left|(\tilde{g}_t^2\tilde{\eta}_t^2 - g_t^2\eta_t^2)(\theta)\right| \leq \frac{1}{\underline{h}}\Big|(2y_t^2 g_t - 6y_t g_t^2 + 4g_t^3)(\tilde{g}_t - g_t) + (y_t^2 - 6y_t g_t + 6g_t^2)(\tilde{g}_t - g_t)^2 + 2(2g_t - y_t)(\tilde{g}_t - g_t)^3 + (\tilde{g}_t - g_t)^4\Big|(\theta) + \frac{1}{\underline{h}^2}\left|g_t^2\epsilon_t^2(\tilde{h}_t - h_t)\right|(\theta).$$

Thus, similarly to (19), we can have

$$\left|\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n) - \frac{1}{n}\sum_{t=1}^{n}\left(g_t^2\eta_t^2\right)(\hat{\theta}_n)\right| = o_{\mathrm{P}}(1).$$

By (M1), (A3) and (A6), one can readily check that for any neighborhood $N(\theta_0)$,

$$\mathbb{E}\sup_{\theta\in N(\theta_0)}\left|\left(g_t^2\eta_t^2\right)(\theta)\right| = \mathbb{E}\sup_{\theta\in N(\theta_0)}\left|\frac{g_t^2(y_t - g_t)^2}{h_t}(\theta)\right| < \infty,$$

and thus, similarly to (23), we have

$$\left|\frac{1}{n}\sum_{t=1}^{n}\left(g_t^2\eta_t^2\right)(\hat{\theta}_n) - \frac{1}{n}\sum_{t=1}^{n}\left(g_t^2\eta_t^2\right)(\theta_0)\right| = o_{\mathrm{P}}(1).$$

Using these, we establish (26).

Next, we show that

$$\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t\tilde{\eta}_t^3\right)(\hat{\theta}_n) \xrightarrow{\mathrm{P}} \mathbb{E}\left(g_t\eta_t^3\right)(\theta_0). \tag{27}$$

Note that

$$\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t\tilde{\eta}_t^3\right)(\hat{\theta}_n) = \left(\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t\tilde{\eta}_t^3\right)(\hat{\theta}_n) - \frac{1}{n}\sum_{t=1}^{n}\left(g_t\eta_t^3\right)(\hat{\theta}_n)\right) + \left(\frac{1}{n}\sum_{t=1}^{n}(g_t\eta_t^3)(\hat{\theta}_n) - \frac{1}{n}\sum_{t=1}^{n}(g_t\eta_t^3)(\theta_0)\right) + \frac{1}{n}\sum_{t=1}^{n}\left(g_t\eta_t^3\right)(\theta_0),$$

$$\left|(\tilde{g}_t\tilde{\eta}_t^3 - g_t\eta_t^3)(\theta)\right| \leq \frac{1}{\underline{h}^{3/2}}\left|\left(y_t^3 - 6y_t^2 g_t + 9y_t g_t^2 - 4g_t^3\right)(\tilde{g}_t - g_t) - \left(3y_t - 9y_t g_t + 6g_t^2\right) \times(\tilde{g}_t - g_t)^2 + (3y_t - 4g_t)(\tilde{g}_t - g_t)^3 - (\tilde{g}_t - g_t)^4\right|(\theta) + \frac{|g_t\eta_t^3|}{\underline{h}^{9/2}}\left|3h_t^2(\tilde{h}_t - h_t) + 3h_t(\tilde{h}_t - h_t)^2 + (\tilde{h}_t - h_t)^3\right|(\theta)$$

and owing to (M1), (A3) and (A6), $\mathbb{E}\sup_{\theta\in N(\theta_0)}|(g_t\eta_t^3)(\theta)| = \mathbb{E}\sup_{\theta\in N(\theta_0)}|\frac{g_t(y_t-g_t)^3}{h_t^{3/2}}(\theta)| < \infty$. Then, similarly to (26), we get (27).

Finally, we show that

$$\frac{1}{n}\sum_{t=1}^{n}\tilde{\eta}_t^4(\hat{\theta}_n) \xrightarrow{\text{P}} \mathbb{E}\eta_t^4. \tag{28}$$

For this, note that

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\tilde{\eta}_t^4(\hat{\theta}_n) &= \left(\frac{1}{n}\sum_{t=1}^{n}\tilde{\eta}_t^4(\hat{\theta}_n) - \frac{1}{n}\sum_{t=1}^{n}\eta_t^4(\hat{\theta}_n)\right)\\ &\quad + \left(\frac{1}{n}\sum_{t=1}^{n}\eta_t^4(\hat{\theta}_n) - \frac{1}{n}\sum_{t=1}^{n}\eta_t^4(\theta_0)\right) + \frac{1}{n}\sum_{t=1}^{n}\eta_t^4(\theta_0),\\ \left(\tilde{\eta}_t^4 - \eta_t^4\right)(\theta) &\le \frac{1}{\underline{h}^2}\Big| -4(y_t-g_t)^3(\tilde{g}_t-g_t) + 6(y_t-g_t)^2(\tilde{g}_t-g_t)^2 - 4(y_t-g_t)(\tilde{g}_t-g_t)^3\\ &\quad + (\tilde{g}_t-g_t)^4\Big|(\theta) + \frac{1}{\underline{h}^4}\Big|2h_t\epsilon_t^4(\tilde{h}_t-h_t) - \epsilon_t^4(\tilde{h}_t-h_t)^2\Big|(\theta),\end{aligned}$$

and owing to (M1), (A3) and (A6), $\mathbb{E}\sup_{\theta\in N(\theta_0)}|\eta_t^4(\theta)| = \mathbb{E}\sup_{\theta\in N(\theta_0)}|\frac{(y_t-g_t)^4}{h_t^2}(\theta)| < \infty$. Then, (28) can be yielded similarly to (26). Combining (26)–(28), we establish the lemma. □

Lemma 4 *Suppose that* (M3), (M4), (A1), (A2), (A4) *and* (A5) *hold. Then, we have that for* $t \le k_0$,

$$\begin{aligned}\sup_{\theta\in\Theta}|\tilde{g}_t(\mu) - g_{1,t}(\mu)| &\le V_g(y_{1,t-1},\ldots)\cdot\rho^t,\\ \sup_{\theta\in\Theta}|\tilde{h}_t(\mu) - h_{1,t}(\mu)| &\le V_h(y_{1,t-1},\ldots)\cdot\rho^t,\quad a.s.\end{aligned}$$

and

$$\begin{aligned}\sup_{\theta\in\Theta}|\partial_\theta\tilde{g}_t(\mu) - \partial_\theta g_{1,t}(\mu)| &\le V_{dg}(y_{1,t-1}\ldots)\cdot\rho^t,\\ \sup_{\theta\in\Theta}|\partial_\theta\tilde{h}_t(\mu) - \partial_\theta h_{1,t}(\mu)| &\le V_{dh}(y_{1,t-1},\ldots)\cdot\rho^t\quad a.s.;\end{aligned}$$

and further, for $t > k_0$,

$$\begin{aligned}\sup_{\theta\in\Theta}|\tilde{g}_t(\mu) - g_{2,t}(\mu)| &\le V_g(y_{2,t-1},\ldots)\cdot\rho^{t-k_0},\\ \sup_{\theta\in\Theta}|\tilde{h}_t(\mu) - h_{2,t}(\mu)| &\le V_h(y_{2,t-1},\ldots)\cdot\rho^{t-k_0}\quad a.s.\end{aligned}$$

and

$$\begin{aligned}\sup_{\theta\in\Theta}|\partial_\theta\tilde{g}_t(\mu) - \partial_\theta g_{1,t}(\mu)| &\le V_{dg}(y_{2,t-1}\ldots)\cdot\rho^{t-k_0},\\ \sup_{\theta\in\Theta}|\partial_\theta\tilde{h}_t(\mu) - \partial_\theta h_{1,t}(\mu)| &\le V_{dh}(y_{2,t-1},\ldots)\cdot\rho^{t-k_0}\quad a.s.\end{aligned}$$

Proof Due to (A2), for some $X' \in L^2(R^\infty)$ and $t \leq k_0$,

$$\begin{aligned}\sup_{\theta\in\Theta} |\tilde{g}_t(\mu) - g_{1,t}(\mu)| &= \sup_{\theta\in\Theta} |g(y_{1,t-1},\ldots,y_{1,1},X';\mu) - g(y_{1,t-1},y_{1,t-2},\ldots;\mu)| \\ &\leq V_g(y_{1,t-1},\ldots)\cdot\rho^t \;\; a.s., \end{aligned} \tag{29}$$

whereas for $t > k_0$,

$$\begin{aligned}\sup_{\theta\in\Theta} |\tilde{g}_t(\mu) - g_{2,t}(\mu)| &= \sup_{\theta\in\Theta} \Big| g(y_{2,t-1},\ldots,y_{2,k_0+1},y_{1,k_0},\ldots,y_{1,1},X';\mu) \\ &\quad - g(y_{2,t-1},y_{2,t-2},\ldots;\mu)\Big| \\ &\leq V_g(y_{2,t-1},\ldots)\cdot\rho^{t-k_0} \;\; a.s. \end{aligned} \tag{30}$$

The rest part of the lemma can be proven similarly to (29) and (30). This completes the proof. □

Lemma 5 *Suppose that* (M2)–(M4), (A1)–(A6), (C1)(i) *and* (C2)(i) *hold. Then, as* $n \to \infty$,

$$\hat{\Sigma}_n \xrightarrow{P} \tau_0\Sigma_1 + (1-\tau_0)\Sigma_2,$$

where $\Sigma_i = Var\Big((g_{i,t}\eta_{i,t})(\tilde{\theta}_0), \eta_{i,t}^2(\tilde{\theta}_0)\Big)$ *for* $i = 1, 2$.

Proof We first show that

$$\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n) \xrightarrow{P} \tau_0\mathbb{E}\left(g_{1,t}^2\eta_{1,t}^2\right)\left(\tilde{\theta}_0\right) + (1-\tau_0)\mathbb{E}\left(g_{2,t}^2\eta_{2,t}^2\right)(\tilde{\theta}_0). \tag{31}$$

Note that

$$\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n) = \frac{1}{n}\sum_{t=1}^{k_0}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n) + \frac{1}{n}\sum_{t=k_0+1}^{n}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n).$$

Owing to Lemma 4, similarly to (26), we have

$$\frac{1}{n}\sum_{t=1}^{k_0}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n) \xrightarrow{P} \tau_0\mathbb{E}\left(g_{1,t}^2\eta_{1,t}^2\right)(\tilde{\theta}_0) \tag{32}$$

and

$$\frac{1}{n}\sum_{t=k_0+1}^{n}\left(\tilde{g}_t^2\tilde{\eta}_t^2\right)(\hat{\theta}_n) \xrightarrow{P} (1-\tau_0)\mathbb{E}\left(g_{2,t}^2\eta_{2,t}^2\right)(\tilde{\theta}_0). \tag{33}$$

This implies (31). Similarly, we can easily obtain:

$$\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{g}_t\tilde{\eta}_t^3\right)(\hat{\theta}_n) \xrightarrow{\mathrm{P}} \tau_0\mathbb{E}\left(g_{1,t}\eta_{1,t}^3\right)(\tilde{\theta}_0) + (1-\tau_0)\mathbb{E}\left(g_{2,t}\eta_{2,t}^3\right)(\tilde{\theta}_0) \tag{34}$$

and

$$\frac{1}{n}\sum_{t=1}^{n}\left(\tilde{\eta}_t^4\right)(\hat{\theta}_n) \xrightarrow{\mathrm{P}} \tau_0\mathbb{E}\eta_{1,t}^4\left(\tilde{\theta}_0\right) + (1-\tau_0)\mathbb{E}\eta_{2,t}^4\left(\tilde{\theta}_0\right). \tag{35}$$

This asserts the lemma. □

Proof of Theorem 2 We first show that

$$\sup_{s\in[0,1]}\left|\frac{1}{n}|\hat{T}^L_{n,[ns]}| - L_{\tau_0}(s)\right| = o_{\mathrm{P}}(1), \tag{36}$$

where

$$L_{\tau_0}(s) = \begin{cases} s(1-\tau_0)E_{1,\tilde{\lambda}_0}, & s < \tau_0, \\ \tau_0(1-\tau_0)E_{1,\tilde{\lambda}_0}, & s = \tau_0, \\ \tau_0(1-s)E_{1,\tilde{\lambda}_0}, & s > \tau_0. \end{cases}$$

We only deal with the case that $k < k_0$ since the other case can be similarly handled. Note that

$$\begin{aligned}
\hat{T}^L_{n,k} &= \sum_{t=1}^{k}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n) - \frac{k}{n}\sum_{t=1}^{n}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n) \\
&= \frac{n-k}{n}\sum_{t=1}^{k}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n) - \frac{k}{n}\left(\sum_{t=k+1}^{k_0}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n) + \sum_{t=k_0+1}^{n}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n)\right) \\
&= \frac{k(n-k)}{n}\frac{1}{k}\sum_{t=1}^{k}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n) - \frac{k}{n}\sum_{t=k+1}^{k_0}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n) \\
&\quad - \frac{k(n-k_0)}{n}\frac{1}{n-k_0}\sum_{t=k_0+1}^{n}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n).
\end{aligned}$$

Similarly to (32) and (33), we have that for $k < k_0$,

$$\frac{1}{k}\sum_{t=1}^{k}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n) \xrightarrow{\mathrm{P}} \mathbb{E}(g_{1,t}\eta_{1,t})(\tilde{\theta}_0),\ \frac{1}{k_0-k}\sum_{t=k+1}^{k_0}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n) \xrightarrow{\mathrm{P}} \mathbb{E}\left(g_{1,t}\eta_{1,t}\right)(\tilde{\theta}_0)$$

and

$$\frac{1}{n-k_0}\sum_{t=k_0+1}^{n}(\tilde{g}_t\tilde{\eta}_t)(\hat{\theta}_n)\xrightarrow{\text{P}}\mathbb{E}(g_{2,t}\eta_{2,t})(\tilde{\theta}_0).$$

Hence, due to (C1), for $0 \le s < t_0$ we can find a constant $E_{1,\tilde{\lambda}_0}$ such that

$$\frac{1}{n}|\hat{T}^L_{n,[ns]}| = s(1-\tau_0)E_{1,\tilde{\lambda}_0} + o_{\text{P}}(1),$$

which in turn implies (36). Because $\hat{\kappa}_n^2 = O_{\text{P}}(1)$ (cf. Lemma 5), the lemma can be established by the arguments similar to those in the proofs of Theorem 5 and Corollary 1 of Kirch and Kamgaing (2012). □

Proof of Theorem 3 The theorem follows from (35) and the arguments similar to those in the proof of Theorem 2. □

Acknowledgements We would like to thank the Editor, an AE and the two referees for their careful reading and valuable comments that improve the quality of the paper.

References

An, H. Z., Huang, F. C. (1996). The geometrical ergodicity of nonlinear autoregressive models. *Statistica Sinica*, *6*, 943–956.

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

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

Brown, R. L., Durbin, J., Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of Royal Statistical Society Series B (Methodological)*, *37*(2), 149–192.

Chen, M., An, H. Z. (1998). A note on the stationarity and the existence of moments of the GARCH model. *Statistica Sinica*, *8*, 505–510.

de Pooter, M., van Dijk, D. (2004). *Testing for changes in volatility in heteroskedastic time series-a further examination* (No. EI 2004-38).

Francq, C., Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, *10*(4), 605–637.

Franke, J., Kirch, C., Kamgaing, J. T. (2012). Changepoints in times series of counts. *Journal of Time Series Analysis*, *33*(5), 757–770.

Gombay, E. (2008). Change detection in autoregressive time series. *Journal of Multivariate Analysis*, *99*(3), 451–464.

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.

Kang, J., Lee, S. (2014). Parameter change test for poisson autoregressive models. *Scandinavian Journal of Statistics*, *41*(4), 1136–1152.

Kiefer, J. (1959). K-sample analogues of the Kolmogorov–Smirnov and Cramer-V Mises tests. *The Annals of Mathematical Statistics*, *30*(2), 420–447.

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*(2), 445–462.

Kirch, C., Kamgaing, J. T. (2012). Testing for parameter stability in nonlinear autoregressive models. *Journal of Time Series Analysis*, *33*(3), 365–385.

Kokoszka, P., Leipus, R. (1999). Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal*, *39*(2), 182–195.

Kulperger, R., Yu, H. (2005). High moment partial sum processes of residuals in GARCH models and their applications. *The Annals of Statistics*, *33*, 2395–2422.

Lee, J., Lee, S. (2015). Parameter change test for nonlinear time series models with GARCH type errors. *Journal of Korean Mathematical Society*, *52*, 503–522.

Lee, S., Na, O. (2005). Test for parameter change in stochastic processes based on conditional least-squares estimator. *Journal of Multivariate Analysis*, *93*, 375–393.

Lee, S., Oh, H. (2016). Parameter change test for autoregressive conditional duration models. *Annals of the Institute of Statistical Mathematics*, *68*(3), 621–637.

Lee, S., Song, J. (2008). Test for parameter change in ARMA models with GARCH innovations. *Statistics and Probability Letters*, *78*, 1990–1998.

Lee, S., Ha, J., Na, O., Na, S. (2003). The cusum test for parameter change in time series models. *Scandinavian Journal of Statistics*, *30*(4), 781–796.

Lee, S., Tokutsu, Y., Maekawa, K. (2004). The cusum test for parameter change in regression models with ARCH errors. *Journal of the Japan Statistical Society*, *34*(2), 173–188.

Meitz, M., Saikkonen, P. (2011). Parameter estimation in nonlinear AR-GARCH models. *Econometric Theory*, *27*(6), 1236–1278.

Oh, H., Lee, S. (2017a). On score vector- and residual-based cusum tests in ARMA-GARCH models. *Statistical Methods and Applications*, *1*, 1–22. https://doi.org/10.1007/s10260-017-0408-9.

Oh, H., Lee, S. (2017b). Bootstrap parameter change test for location-scale time series models with heteroscedasticity. **(Submitted for publication)**.

Page, E. S. (1955). A test for change in a parameter occurring at an unknown point. *Biometrika*, *42*(3/4), 523–527.

Straumann, D., Mikosch, T. (2006). Quasi maximum likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics*, *34*(5), 2449–2495.