Ann Inst Stat Math (2016) 68:621–637
DOI 10.1007/s10463-015-0541-x

Parameter change test for autoregressive conditional duration models

Sangyeol Lee[1] · **Haejune Oh**[1]

Received: 2 May 2014 / Revised: 30 December 2014 / Published online: 9 September 2015

Abstract In this study, we consider the parameter change test in nonlinear autoregressive conditional duration models. Particularly, we use the cumulative sum test based on parameter estimates and verify that its limiting null distribution is the supremum of a Brownian bridge. A simulation study and real data analysis are provided for illustration.

Keywords Nonlinear ACD models · Parameter change test · Cusum test · Brownian bridge

1 Introduction

Since Page (1955), testing for a parameter change has played an important role in economics, engineering and medicine, and a vast number of articles exist in various research areas: see Csörgő and Horváth (1997). The change-point problem has drawn much attention from the researchers in financial time series analysis because time series often suffer from structural changes owing to policy changes and critical social events and ignoring it can lead to a false conclusion. The cumulative sum (cusum) test has been broadly used to detect parameter changes since it is easy to implement in many applications. Inclán and Tiao (1994) designed a variance change test. Later, their method has been extended to various time series models such as ARMA–GARCH,

Electronic supplementary material The online version of this article (doi:10.1007/s10463-015-0541-x) contains supplementary material, which is available to authorized users.

✉ Sangyeol Lee
jpgrslee@yahoo.com

1 Department of Statistics, Seoul National University, Seoul 151-747, Korea

ARCH regression, Poisson AR, multivariate GARCH models, and diffusion processes. We refer to Lee et al. (2003, 2006), Lee and Song (2008), Kim and Lee (2009), Na et al. (2012, 2013), and Kang and Lee (2014). However, these articles do not cover all important time series models. Particularly, to our knowledge, the cusum test has not yet been investigated for autoregressive conditional duration (ACD) models. Motivated by this, we are led to study the cusum test for the ACD models.

Originally, Engle and Russell (1997, 1998) proposed ACD models to analyze irregularly spaced high-frequency data on durations between successive financial market trades. Their method is well fitted to modeling transaction data with irregular time intervals and studying the dynamics of the corresponding durations. Engle and Russell (1997) fitted an ACD model with a linear specification to IBM transactions data, but later, Dufour and Engle (2000) pointed out that the linear specification could be too restrictive in practice. Since then, many alternative nonlinear ACD models have been proposed by many authors. For example, Bauwens and Giot (2000) proposed the logarithmic ACD (log-ACD) models, Dufour and Engle (2000) introduced the Box-Cox ACD (BCACD) and the exponential ACD (EXPACD) models, and Zhang et al. (2001) proposed the threshold ACD (TACD) models. More recent developments can be found in Bauwens and Giot (2003) who proposed asymmetric ACD models, Bauwens and Veredas (2004) who proposed stochastic conditional duration (SCD) models, Fernandes and Grammig (2006) who proposed augmented ACD (AACD) models, and Meitz and Terasvirta (2006) who propose smooth-transition threshold (ST-ACD) models. In this study, we consider the cusum test within the framework of these nonlinear ACD models. Since our cusum method is based on the quasi-maximum likelihood estimator (QMLE) from ACD models and heavily relies on its asymptotic properties, we investigate the strong consistency and asymptotic normality of the QMLE. The result is then used to derive the limiting null distribution of the cusum test.

The remainder of this paper is organized as follows. In Sect. 2.1, we present the asymptotic properties of the QMLE in nonlinear ACD models. In Sect. 2.2, we present the cusum test in nonlinear ACD models and show that under regularity conditions, the cusum test based on the QMLE converges weakly to the supremum (sup) of a Brownian bridge. In Sects. 3 and 4, we perform a simulation study and real data analysis, focusing on the ACD and log-ACD models. In Sect. 5, we provide concluding remarks. In the Appendix, we provide the proofs of the theorems in Sect. 2.

2 Cusum test for nonlinear ACD models

2.1 Asymptotics for nonlinear ACD models

Let us consider the nonlinear ACD(p, q) model:

$$\begin{cases} x_i &= \psi_i \epsilon_i, \\ \psi_i &= g_{\boldsymbol{\theta}}(x_{i-1}, x_{i-2}, \ldots, x_{i-p}; \psi_{i-1}, \psi_{i-2}, \ldots, \psi_{i-q}), \end{cases} \tag{1}$$

where $\{g_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ denotes a parametric family of non-negative functions on $[0, \infty)^p \times [0, \infty)^q$, and $\{\epsilon_i\}$ is a sequence of independently and identically distributed (iid) positive random variables with $\mathbb{E}\epsilon_0 = 1$. We also assume that ψ_i is non-negative and $\mathcal{F}_{i-1}$-measurable, where $\mathcal{F}_i = \sigma(\epsilon_k; k \le i)$. In practice, x_i is the duration between two consecutive events, that is, $t_i - t_{i-1}$, where t_i is the time that the ith event (for instance, trade, quote, price change, etc.) occurs. The ψ_i denotes the conditional expected value given past observations and ϵ_i are iid error terms. Basically, the structure of the ACD models is similar in spirit to that of the GARCH models.

We express $\psi_i = g_{\boldsymbol{\theta}}(\mathbf{X}_{i-1}, \Psi_{i-1})$, where $\mathbf{X}_i = (x_i, \ldots, x_{i-p+1})^{\mathrm{T}}$ and $\Psi_i = (\psi_i, \ldots, \psi_{i-q+1})^{\mathrm{T}}$. Let K be a compact subset of Θ. For any initial value $\boldsymbol{\zeta}_0 \in [0, \infty)^q$, we recursively define the following random vector functions $\hat{\Psi}_i$ on K:

$$\hat{\Psi}_i = \begin{cases} \boldsymbol{\zeta}_0, & i = 0, \\ \phi_{i-1}(\hat{\Psi}_{i-1}), & i \ge 1, \end{cases}$$

where the random maps $\phi_i : \mathbb{C}(K, [0, \infty)^q) \to \mathbb{C}(K, [0, \infty)^q)$ are defined by

$$[\phi_i(\boldsymbol{a})](\boldsymbol{\theta}) = (g_{\boldsymbol{\theta}}(\mathbf{X}_i, \boldsymbol{a}(\boldsymbol{\theta})), a_1(\boldsymbol{\theta}), a_2(\boldsymbol{\theta}), \ldots, a_{q-1}(\boldsymbol{\theta}))^{\mathrm{T}}$$

with $\boldsymbol{a} = (a_1, \ldots, a_q)^{\mathrm{T}} \in \mathbb{C}(K, [0, \infty)^q)$, the space of continuous $[0, \infty)^q$-valued functions equipped with the sup-norm $\|\boldsymbol{a}\|_K = \sup_{s \in K} |\boldsymbol{a}(s)|$. We can regard $\hat{\Psi}_i(\boldsymbol{\theta}) = (\hat{\psi}_i(\boldsymbol{\theta}), \ldots, \hat{\psi}_{i-q+1}(\boldsymbol{\theta}))^{\mathrm{T}}$ an "estimate" of the Ψ_i under the parameter hypothesis $\boldsymbol{\theta}$.

Suppose that $x_{-p+1}, \ldots, x_0, x_1, \ldots, x_n$ are generated from Model (1) with the true parameter $\boldsymbol{\theta}_0$, wherein $\hat{\psi}_i$, $1 \le i \le n$, are well defined. We define the conditional exponential likelihood by

$$\hat{L}_n(\boldsymbol{\theta}) = -\sum_{i=1}^{n} \left(\frac{x_i}{\hat{\psi}_i(\boldsymbol{\theta})} + \log \hat{\psi}_i(\boldsymbol{\theta}) \right) := -\sum_{i=1}^{n} \hat{l}_i(\boldsymbol{\theta})$$

and the exponential QMLE by

$$\hat{\boldsymbol{\theta}}_n := \underset{\boldsymbol{\theta} \in K}{\operatorname{argmax}}\, \hat{L}_n(\boldsymbol{\theta}).$$

Further, we set

$$L_n(\boldsymbol{\theta}) = -\sum_{i=1}^{n} \left(\frac{x_i}{\psi_i(\boldsymbol{\theta})} + \log \psi_i(\boldsymbol{\theta}) \right) := \sum_{i=1}^{n} l_i(\boldsymbol{\theta})$$

and

$$\tilde{\boldsymbol{\theta}}_n := \underset{\boldsymbol{\theta} \in K}{\operatorname{argmax}}\, L_n(\boldsymbol{\theta}).$$

In the following, we denote $\phi_i^{(r)} = \phi_i \circ \phi_{i-1} \circ \cdots \circ \phi_{i-r+1}$, $r \ge 1$, for a sequence of random mappings $\{\phi_i\}$ defined on $\mathbb{C}(K, [0, \infty)^q)$.

The norm of a matrix $A = (a_{ij}) \in \mathbf{R}^{n\times n}$ for any $n \in \mathbb{N}$ is defined as $\|A\| = (\sum a_{ij}^2)^{1/2}$, and the norm of a continuous matrix-valued function A on any compact set $M \in \mathbf{R}^d$ is defined as $\|A\|_M = \sup_{s\in M} \|A(s)\|$. Further, we set $\Lambda(\phi_i) = \sup_{x,y\in\mathbb{C}(M,[0,\infty)^q),x\neq y} \left(\frac{\|\phi_i(x)-\phi_i(y)\|_M}{\|x-y\|_M}\right)$ and use the notation $v_n \xrightarrow{\text{e.a.s.}} 0$ (v_n converges to zero exponential fast a.s.) as $n \to \infty$ when there exist $\gamma > 1$ with $\gamma^n \|v_n\| \xrightarrow{\text{a.s.}} 0$. We define $\log^+(x) = \log x$ if $x > 1$ and 0 otherwise.

To verify the strong consistency of the QMLE, we assume the following conditions:

(C.1) Model (1.1) with $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ admits a unique stationary ergodic solution $\{(x_i, \psi_{i,0})\}$ with $\mathbb{E}(\log^+ \psi_{0,0}) < \infty$.

(C.2) For every fixed $\mathbf{x} \in [0,\infty)^p$, the map $(\boldsymbol{\theta}, \boldsymbol{\psi}) \mapsto g_{\boldsymbol{\theta}}(\mathbf{x}, \boldsymbol{\psi})$ is continuous; $\boldsymbol{\theta}_0$ is an interior point of K; $\mathbb{E}(\log^+ \|\phi_0(\boldsymbol{\zeta}_0)\|_K) < \infty$; $\mathbb{E}[\log^+ \Lambda(\phi_0)] < \infty$; there exists an integer $r \geq 1$ such that $\mathbb{E}[\log \Lambda(\phi_0^{(r)})] < 0$, where $\mathbb{E}$ denotes the expectation under the true parameter $\boldsymbol{\theta}_0$.

(C.3) The class of functions $\{g_{\boldsymbol{\theta}} | \boldsymbol{\theta} \in K\}$ is uniformly bounded from below, that is, there exists a constant $\underline{g} > 0$ such that $g_{\boldsymbol{\theta}}(\mathbf{x}, \boldsymbol{\psi}) \geq \underline{g}$ for all $(\mathbf{x}, \boldsymbol{\psi}) \in [0,\infty)^p \times [0,\infty)^q$ and $\boldsymbol{\theta} \in K$.

(C.4) For all $\boldsymbol{\theta} \in K$,

$$\psi_0(\boldsymbol{\theta}) \equiv \psi_{0,0} \quad \text{a.s. if} \quad \text{and} \quad \text{only if} \quad \boldsymbol{\theta} = \boldsymbol{\theta}_0.$$

These conditions are found in Straumann and Mikosch (2006). In fact, the proof of the consistency is similar to that of Theorem 4.1 of Straumann and Mikosch (2006). The detailed proof is provided in the supplementary material.

Theorem 1 *Under conditions* (C.1)–(C.4), *we have*

$$\hat{\boldsymbol{\theta}}_n \xrightarrow{\text{a.s.}} \boldsymbol{\theta}_0 \quad as \quad n \to \infty.$$

Now, we turn our attention to the asymptotic normality of the QMLE. For this, we introduce a map $\varphi_i : K \times [0,\infty)^q \to [0,\infty)^q$ such as $\varphi_i(\boldsymbol{\theta}, \boldsymbol{y}) \mapsto (g_{\boldsymbol{\theta}}(\mathbf{X}_i, \boldsymbol{y}), y_1, \ldots, y_{q-1})$ with $\boldsymbol{y} = (y_1, \ldots, y_q)^{\mathrm{T}}$: here, $[\phi_i(\boldsymbol{a})](\boldsymbol{\theta}) = \varphi_i(\boldsymbol{\theta}, \boldsymbol{a}(\boldsymbol{\theta}))$ for every $\boldsymbol{a} \in \mathbb{C}(K, [0,\infty)^q)$. In what follows, we use the notation $[\boldsymbol{a}]_{(j)}$ and $[A]_{(j,k)}$ to denote the jth component of vector $\boldsymbol{a}$ and the (j,k)th entry of matrix A, respectively. Further, for any real-valued function h defined on a subset of R^l, $l \geq 1$, we denote $\partial_k h(x_1,\ldots,x_l) = \frac{\partial h(x_1,\ldots,x_l)}{\partial x_k}$ and $\partial^2_{k_1,k_2} h(x_1,\ldots,x_l) = \frac{\partial^2 h(x_1,\ldots,x_l)}{\partial x_{k_1}\partial x_{k_2}}$; for vector function $\mathbf{f}$ with its jth component f_j, $\mathbf{f}'$ denotes the matrix whose (j,k)th entry is $\partial_k f_j$.

Since $\{\hat{\Psi}_i\}$ is a solution of the stochastic recurrence equation (SRE) $\boldsymbol{a}_{i+1} = \phi_i(\boldsymbol{a}_i)$ on $\mathbb{C}(K, [0,\infty)^q)$, as in Straumann and Mikosch (2006), it can be seen that the differentiation with respect to $\boldsymbol{\theta}$ on both sides of $\hat{\Psi}_{i+1}(\boldsymbol{\theta}) = \phi_i(\hat{\Psi}_i(\boldsymbol{\theta})) = \varphi_i(\boldsymbol{\theta}, \hat{\Psi}_i(\boldsymbol{\theta}))$ leads to the SRE $\hat{\Psi}'_{i+1} = \hat{\dot{\phi}}_i(\hat{\Psi}'_i)$ on $\mathbb{C}(K, \mathbf{R}^{q\times d})$, namely,

$$[\partial_k \hat{\Psi}_{i+1}(\boldsymbol{\theta})]_{(j)} = [\partial_k \varphi_i(\boldsymbol{\theta}, \hat{\Psi}_i(\boldsymbol{\theta}))]_{(j)} + \sum_{l=1}^{q} [\partial_{(d+l)} \varphi_i(\boldsymbol{\theta}, \hat{\Psi}_i(\boldsymbol{\theta}))]_{(j)} [\partial_k \hat{\Psi}_i(\boldsymbol{\theta})]_{(l)} \quad (2)$$

for $j \in \{1, \ldots, q\}$ and $k \in \{1, \ldots, d\}$. Then, the replacement of $\hat{\Psi}_i$ and $\hat{\Psi}_i'$ by Ψ_i and $\mathbf{D}_i$ in (2) leads to the SRE $\mathbf{D}_{i+1} = \dot{\phi}_i(\mathbf{D}_i)$ on $\mathbb{C}(K, \mathbf{R}^{q\times d})$, namely,

$$[\mathbf{D}_{i+1}(\boldsymbol{\theta})]_{(j,k)} = [\partial_k \varphi_i(\boldsymbol{\theta}, \Psi_i(\boldsymbol{\theta}))]_{(j)} + \sum_{l=1}^{q} [\partial_{(d+l)} \varphi_i(\boldsymbol{\theta}, \Psi_i(\boldsymbol{\theta}))]_{(j)} [\mathbf{D}_i(\boldsymbol{\theta})]_{(l,k)}. \quad (3)$$

Further, in the same spirit of Straumann and Mikosch (2006), we can also express $\hat{\Psi}_{i+1}'' = \hat{\ddot{\phi}}_i(\hat{\Psi}_i'')$ by differentiating both sides of (2) with respect to $\boldsymbol{\theta}$ and obtain the SRE $\mathbf{E}_{i+1} = \ddot{\phi}_i(\mathbf{E}_i)$ on $\mathbb{C}(K, \mathbf{R}^{q\times d^2})$ similarly to the argument $\mathbf{D}_{i+1} = \dot{\phi}_i(\mathbf{D}_i)$.

Below, we present the conditions to guarantee the existence of the first and second derivatives of ψ_i:

(D.1) (C.1) and (C.2) hold with $K \subset \mathbf{R}^d$; K coincides with the closure of its interior; and for every fixed $\mathbf{x} \in \mathbf{R}^p$, the function $(\boldsymbol{\theta}, \boldsymbol{\psi}) \mapsto g_{\boldsymbol{\theta}}(\mathbf{x}, \boldsymbol{\psi})$ on $K \times [0, \infty)^q$ is continuously differentiable.

(D.2) For all $j \in \{1, \ldots, q\}$ and $k \in \{1, \ldots, d+q\}$,

$$\mathbb{E}\left[\log^+\left(\sup_{\boldsymbol{\theta}\in K} \left|\partial_k[\varphi_0(\boldsymbol{\theta}, \Psi_0(\boldsymbol{\theta}))]_{(j)}\right|\right)\right] < \infty,$$

and furthermore, there exists a stationary sequence $\{\bar{C}_1(i)\}$ with $\mathbb{E}[\log^+ \bar{C}_1(0)] < \infty$ and $\kappa \in (0, 1]$ such that

$$\sup_{\boldsymbol{\theta}\in K} \left|\partial_k[\varphi_i(\boldsymbol{\theta}, \boldsymbol{a})]_{(j)} - \partial_k[\varphi_i(\boldsymbol{\theta}, \tilde{\boldsymbol{a}})]_{(j)}\right| \le \bar{C}_1(i)|\boldsymbol{a} - \tilde{\boldsymbol{a}}|^{\kappa}, \quad \boldsymbol{a}, \tilde{\boldsymbol{a}} \in [0, \infty)^q$$

for every $j \in \{1, \ldots, q\}$ and $k \in \{1, \ldots, d+q\}$.

(D.3) For every fixed $\mathbf{x} \in \mathbf{R}^p$, the function $(\boldsymbol{\theta}, \boldsymbol{\psi}) \mapsto g_{\boldsymbol{\theta}}(\mathbf{x}, \boldsymbol{\psi})$ on $K \times [0, \infty)^q$ is twice continuously differentiable; for all $j \in \{1, \ldots, q\}$ and $k_1, k_2 \in \{1, \ldots, d+q\}$,

$$\mathbb{E}\left[\log^+\left(\sup_{\boldsymbol{\theta}\in K} \left|\partial^2_{k_1,k_2}[\varphi_0(\boldsymbol{\theta}, \Psi_0(\boldsymbol{\theta}))]_{(j)}\right|\right)\right] < \infty;$$

the sequence of the first derivative $\{\Psi_i'\}$ satisfies $\mathbb{E}(\log^+ \|\Psi_0'\|_K) < \infty$; there exists a stationary sequence $\{\bar{C}_2(i)\}$ with $\mathbb{E}[\log^+ \bar{C}_2(0)] < \infty$ and $\bar{\kappa} \in (0, 1]$ such that

$$\sup_{\boldsymbol{\theta}\in K} \left|\partial^2_{k_1,k_2}[\varphi_i(\boldsymbol{\theta}, \boldsymbol{a})]_{(j)} - \partial^2_{k_1,k_2}[\varphi_i(\boldsymbol{\theta}, \tilde{\boldsymbol{a}})]_{(j)}\right| \le \bar{C}_2(i)|\boldsymbol{a} - \tilde{\boldsymbol{a}}|^{\bar{\kappa}}, \quad \boldsymbol{a}, \tilde{\boldsymbol{a}} \in [0, \infty)^q$$

for every $j \in \{1, \ldots, q\}$ and $k_1, k_2 \in \{1, \ldots, d+q\}$.

Remark 1 According to Straumann and Mikosch (2006), provided that (D.1) and (D.2) hold, SRE (3) has a unique ergodic stationary solution $\{\mathbf{D}_i\}$. In addition, $\mathbf{D}_i$ is $\mathcal{F}_{i-1}$-measurable and $\|\hat{\Psi}_i' - \mathbf{D}_i\|_K \xrightarrow{\text{e.a.s.}} 0$ as $i \to \infty$, namely, $\{\mathbf{D}_i\}$ is a stationary approximation of $\{\hat{\Psi}_i'\}$. In addition, the random functions Ψ_i are a.s. continuously differentiable on K, and thus, $\mathbf{D}_i$ becomes the first derivative of Ψ_i: in other words, $\mathbf{D}_i \equiv \Psi_i'$. Moreover, provided that (D.1)–(D.3) hold, the SRE $\mathbf{E}_{i+1} = \ddot{\phi}_i(\mathbf{E}_i)$ has a unique stationary solution $\{\mathbf{E}_i\}$ and $\|\hat{\Psi}_i'' - \mathbf{E}_i\|_K \xrightarrow{\text{e.a.s.}} 0$ as $i \to \infty$. Therefore, $\mathbf{E}_i$ becomes the second derivative of Ψ_i, that is, $\mathbf{E}_i \equiv \Psi_i''$.

To assert the asymptotic normality of the exponential QMLE, we assume the following conditions:

(N.1) (C.1)–(C.4) are fulfilled.
(N.2) (D.1)–(D.3) are fulfilled.
(N.3) The following moment conditions hold:
  (i) $\mathbb{E}\epsilon_0^2 < \infty$,
  (ii) $\mathbb{E}\left(\frac{\|\psi_0'(\boldsymbol{\theta}_0)\|^2}{\psi_{0,0}^2}\right) < \infty$,
  (iii) $\mathbb{E}\|l_0'\|_K < \infty$,
  (iv) $\mathbb{E}\|l_0''\|_K < \infty$.
(N.4) The components of the vector $\frac{\partial g_{\boldsymbol{\theta}}}{\partial \boldsymbol{\theta}}(X_0, \Psi_{0,0})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$ are linearly independent.

(N.2) allows the differentiation of L_n and with probability 1,

$$l_i'(\boldsymbol{\theta}) = -\frac{\psi_i'(\boldsymbol{\theta})}{\psi_i(\boldsymbol{\theta})}\left(1 - \frac{x_i}{\psi_i(\boldsymbol{\theta})}\right),$$
$$l_i''(\boldsymbol{\theta}) = -\frac{1}{\psi_i(\boldsymbol{\theta})^2}\left((\psi_i'(\boldsymbol{\theta})(\psi_i'(\boldsymbol{\theta}))^{\mathrm{T}}\left(2\frac{x_i}{\psi_i(\boldsymbol{\theta})} - 1\right) + \psi_i''(\boldsymbol{\theta})(\psi_i(\boldsymbol{\theta}) - x_i)\right).$$

Remark 2 Owing to Propositions 3.1, 3.12, 6.1, and 6.2 of Straumann and Mikosch (2006), $\{l_i'\}$ and $\{l_i''\}$ are stationary ergodic sequences of random elements with values in $\mathbb{C}(K, \mathbf{R}^d)$ and $\mathbb{C}(K, \mathbf{R}^{d\times d})$, respectively.

Now, we are ready to state the asymptotic normality of the QMLE. Its proof is similar to that of Theorem 7.1 of Straumann and Mikosch (2006) and is provided in the supplementary material.

Theorem 2 *Under conditions* (N.1)–(N.4), $\hat{\boldsymbol{\theta}}_n$ *is asymptotically normal*, *that is*,

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{V}_0) \quad \text{as } n \to \infty,$$

where $\boldsymbol{V}_0 = \mathbb{E}(\epsilon_0^2 - 1)(\mathbb{E}[\psi_0'(\boldsymbol{\theta}_0)(\psi_0'(\boldsymbol{\theta}_0))^{\mathrm{T}}/\psi_{0,0}^2])^{-1}$.

Remark 3 It can be easily seen that ACD(1, 1) model $\psi_i = \omega + \alpha x_{i-1} + \beta\psi_{i-1}$ with the conditions below satisfies (C.1)–(C.4) and (N.1)–(N.4):

(A.1) $\boldsymbol{\theta} = (\omega, \alpha, \beta)$ belongs to a compact set $K \subset (0, \infty) \times [0, \infty) \times [0, 1)$; the true parameter $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0)$ is an interior point of K.
(A.2) $\mathbb{E}\log(\alpha_0\epsilon_0 + \beta_0) < 0$; when $\beta_0 \neq 0$, $\alpha_0 \neq 0$.
(A.3) ϵ_0 has a non-degenerate distribution with $\mathbb{E}\epsilon_0 = 1$ and $\mathbb{E}\epsilon_0^2 < \infty$.

The log-ACD(1, 1) model below could violate (C.2), (N.2) and (N.4):

$$x_i = \psi_i \epsilon_i, \quad \log \psi_i = \omega + \alpha \log x_{i-1} + \beta \log \psi_{i-1}. \tag{4}$$

Nevertheless, Theorems 1 and 2 still hold under some regularity conditions presented below:

Proposition 1 *Suppose that* $\{x_i\}$ *in* (4) *satisfies*

- (L.1) $\boldsymbol{\theta} = (\omega, \alpha, \beta)$ *belongs to a compact set* $K \subset (-\infty, \infty) \times (-\infty, \infty) \times (-1, 1)$; *the true parameter* $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0)$ *is an interior point of* K.
- (L.2) $\alpha_0 > 0$ *and* $0 < \alpha_0 + \beta_0 < 1$.
- (L.3) ϵ_0 *has a non-degenerate distribution with* $\mathbb{E}\epsilon_0 = 1$, $\mathbb{E}\epsilon_0^\nu < \infty$ *and* $\mathbb{E}|\log \epsilon_0|^\nu < \infty$ *for some* $\nu \geq 2$.
- (L.4) *There exists a positive real number* c *such that* $\psi_i(\boldsymbol{\theta}) \geq c$ *for all* $\boldsymbol{\theta} \in K$ *and* i.

Then, $\hat{\boldsymbol{\theta}}_n$ *is strongly consistent and asymptotically normal.*

We provide the proof of Proposition 1 in the supplementary material. The last condition seems necessary to establish Proposition 1. Particularly, it holds if the density of ϵ_i is bounded away from zero. For instance, one can consider $\epsilon_i = \delta_i \vee c$ with positive random variables δ_i and positive (small) constant c. In practice, though, the c can be taken arbitrarily small, so that positive ϵ_i's without such truncation could be considered as errors. In our simulation study (see Sect. 3), exponential ϵ_i's are employed and the cusum test in this case turns out to perform adequately.

2.2 Cusum test

Suppose that given observations $x_{-p+1}, \ldots, x_0, x_1, \ldots, x_n$, we wish to test the following hypotheses:

H_0: The true parameter $\boldsymbol{\theta}$ does not change over $x_{-p+1}, \ldots, x_0, x_1, \ldots, x_n$. vs.
H_1: not H_0.

To perform a test, we employ the cusum test based on the statistic:

$$T_n = \max_{1 \leq k \leq n} \frac{k^2}{n} (\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n)^{\mathrm{T}} (\hat{\boldsymbol{V}}_0^{(n)})^{-1} (\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n), \tag{5}$$

where $(\hat{\boldsymbol{V}}_0^{(n)})^{-1}$ is a consistent estimator of $\boldsymbol{V}_0^{-1}$.

To obtain the limiting null distribution of T_n, we derive a functional central limit theorem for $\hat{\boldsymbol{\theta}}_{[ns]}$, $0 \leq s \leq 1$. Because $\hat{L}'_{[ns]}(\hat{\boldsymbol{\theta}}_{[ns]}) = 0$, by Taylor's expansion theorem,

$$0 = \hat{L}'_{[ns]}(\boldsymbol{\theta}_0) + \int_0^1 \hat{L}''_{[ns]}(\boldsymbol{\theta}_0 + \lambda(\hat{\boldsymbol{\theta}}_{[ns]} - \boldsymbol{\theta}_0)) \mathrm{d}\lambda (\hat{\boldsymbol{\theta}}_{[ns]} - \boldsymbol{\theta}_0)$$

and thus,

$$\frac{\hat{L}'_{[ns]}(\boldsymbol{\theta}_0)}{\sqrt{[ns]}} = \tilde{\mathbf{C}}_{[ns]}\sqrt{[ns]}(\hat{\boldsymbol{\theta}}_{[ns]} - \boldsymbol{\theta}_0),$$

where $\tilde{\mathbf{C}}_{[ns]} = -\frac{1}{[ns]}\int_0^1 \hat{L}''_{[ns]}(\boldsymbol{\theta}_0 + \lambda(\hat{\boldsymbol{\theta}}_{[ns]} - \boldsymbol{\theta}_0))\mathrm{d}\lambda$. Therefore, we can express

$$\mathbf{V}_0^{-\frac{1}{2}}\frac{[ns]}{\sqrt{n}}(\hat{\boldsymbol{\theta}}_{[ns]} - \boldsymbol{\theta}_0) = \mathbf{V}_0^{-\frac{1}{2}}\mathbf{C}_0^{-1}\frac{\hat{L}'_{[ns]}(\boldsymbol{\theta}_0)}{\sqrt{n}} + \mathbf{V}_0^{-\frac{1}{2}}\mathbf{C}_0^{-1}\frac{\sqrt{[ns]}}{\sqrt{n}}\tilde{\Delta}_{[ns]},$$

where $\tilde{\Delta}_{[ns]} = (\mathbf{C}_0 - \tilde{\mathbf{C}}_{[ns]})\sqrt{[ns]}(\hat{\boldsymbol{\theta}}_{[ns]} - \boldsymbol{\theta}_0)$ and $\mathbf{C}_0 = \mathbb{E}[\psi'_0(\boldsymbol{\theta}_0)(\psi'_0(\boldsymbol{\theta}_0))^{\mathrm{T}}/\psi^2_{0,0}]$. Here, $\mathbf{V}_0^{-1/2}\mathbf{C}_0^{-1} = \mathbb{E}[l'_0(\boldsymbol{\theta}_0)(l'_0(\boldsymbol{\theta}_0))^{\mathrm{T}}]^{-1/2}$. Lemma 4 shows that the last term on the right-hand side (RHS) of this equation is asymptotically negligible. Therefore, using Lemma 1, we obtain the following result.

Theorem 3 *Suppose that assumptions* (N.1)–(N.4) *hold with* $\mathbb{E}(\log^+ \|\psi''_0\|_K) < \infty$. *Then, under* H_0, *we have*

$$\mathbf{V}_0^{-\frac{1}{2}}\frac{[ns]}{\sqrt{n}}(\hat{\boldsymbol{\theta}}_{[ns]} - \boldsymbol{\theta}_0) \xrightarrow{w} \mathrm{W}_d(s) \quad in \quad \mathbb{D}([0,1], \mathbf{R}^d),$$

where W_d *is a d-dimensional Wiener process, and thus,*

$$T_n^0 := \max_{1\le k\le n}\frac{k^2}{n}(\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n)^{\mathrm{T}}\mathbf{V}_0^{-1}(\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_n) \xrightarrow{w} \sup_{0\le s\le 1}\|\mathrm{W}_d^0(s)\|^2,$$

where W_d^0 *is a d-dimensional Brownian bridge.*

In practice, we have to replace $\mathbf{V}_0^{-1}$ by its consistent estimator. Using the residuals defined by $\hat{\epsilon}_i^{(n)} = x_i/(\hat{\psi}_i(\hat{\boldsymbol{\theta}}_n))$, we can obtain

$$(\hat{\mathbf{V}}_0^{(n)})^{-1} = \left(\frac{1}{n}\sum_{i=1}^n(\hat{\epsilon}_i^2 - 1)\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n\frac{\hat{\psi}'_i(\hat{\boldsymbol{\theta}}_n)(\hat{\psi}'_i(\hat{\boldsymbol{\theta}}_n))^{\mathrm{T}}}{\hat{\psi}_i(\hat{\boldsymbol{\theta}}_n)^2}\right).$$

To ensure the consistency of $(\hat{\mathbf{V}}_0^{(n)})^{-1}$ (Lemma 5), besides (N.1)–(N.4), a condition like $\mathbb{E}\|\psi'_0/\psi_0\|_K^2 < \infty$ is additionally required: see the Appendix.

Theorem 4 *Let* T_n *be the one defined in* (5). *Suppose that* $\mathbb{E}\|\psi'_0/\psi_0\|_K^2 < \infty$, $\mathbb{E}(\log^+ \|\psi''_0\|_K) < \infty$, *and* (N.1)–(N.4) *hold. Then, under* H_0,

$$T_n \xrightarrow{w} \sup_{0\le s\le 1}\|\mathrm{W}_d^0(s)\|^2.$$

Table 1 $(\omega, \alpha, \beta) = (0.1, 0.2, 0.1)$ in the ACD(1, 1) model

$(\omega', \alpha', \beta')$	$n = 500$	$n = 1000$	$n = 3000$
Size	0.220	0.152	0.128
(0.1, 0.2, 0.5)	1.000	1.000	1.000
(0.1, 0.6, 0.1)	0.994	1.000	1.000
(0.3, 0.2, 0.1)	1.000	1.000	1.000

We reject H_0 if T_n is large. The empirical $(1 - \alpha)$ quantile values for $\sup_{0\le s\le 1} \|\mathrm{W}_p^0(s)\|^2$, $\alpha \in \{0.01, 0.05, 0.1\}$, $p \in \{1, \ldots, 10\}$ can be found in Table 1 of Lee et al. (2003), p. 784.

Particularly, it can be easily seen that the result of Theorem 4 holds for the ACD(1, 1) model satisfying (A.1)–(A.3) in Remark 1. Further, the above theorem still holds for the log-ACD(1, 1) model, the proof of which is provided in the supplementary material.

Proposition 2 *Suppose that the log-ACD*(1, 1) *process* $\{x_i\}$ *in* (4) *satisfies the conditions in Proposition* 1. *Then, under* H_0,

$$T_n \xrightarrow{w} \sup_{0\le s\le 1} \|\mathrm{W}_d^0(s)\|^2.$$

3 Simulation study

In this section, we evaluate the performance of the test statistic T_n through a simulation study. In this simulation study, we perform all tests at the nominal level 0.1. The empirical sizes and powers are calculated as the rejection number of the null hypothesis, under which no model parameters are assumed to experience changes, out of 500 repetitions. To see the performance of T_n, we focus on the ACD(1, 1) and log-ACD(1, 1) models with standard exponential distributed innovations($\{\epsilon_i\} \sim \mathrm{Exp}(1)$). We choose these models because the computational process in these cases is much less time consuming.

We consider the problem of testing the following hypotheses:

H_0: $\boldsymbol{\theta}$ are constant during the observation $i = 1, \ldots, n$. vs.
H_1: $\boldsymbol{\theta}$ change to $\boldsymbol{\theta}'$ at $[n/2]$.

We first evaluate the QMLE $\hat{\theta}_n$ and implement a test with sample sizes $n = 500, 1000, 3000$. The empirical sizes and powers are summarized in Tables 1, 2, 3 and 4, which show that the test has no severe size distortions in most cases, and as anticipated, the empirical size gets closer to the nominal level 0.1 as n increases in all cases. There is a little power loss with sample size $n = 500$, but the power loss is not so significant when the sample size is large enough.

It is quite well known that the cusum test produces less powers when the change point is not located in the middle. To showcase this, we implement the cusum test for the log-ACD model in the same setting as before under the alternatives as follows:

Table 2 $(\omega, \alpha, \beta) = (0.3, 0.1, 0.2)$ in the ACD(1, 1) model

$(\omega', \alpha', \beta')$	$n = 500$	$n = 1000$	$n = 3000$
Size	0.234	0.184	0.104
(0.7, 0.1, 0.2)	1.000	1.000	1.000
(0.3, 0.5, 0.2)	0.998	1.000	1.000
(0.3, 0.1, 0.8)	1.000	1.000	1.000

Table 3 $(\omega, \alpha, \beta) = (0.1, 0.2, 0.1)$ in the log-ACD(1, 1) model

$(\omega', \alpha', \beta')$	$n = 500$	$n = 1000$	$n = 3000$
Size	0.154	0.094	0.104
(0.1, 0.2, 0.5)	0.518	0.714	0.988
(0.1, 0.6, 0.1)	1.000	1.000	1.000
(0.3, 0.2, 0.1)	0.664	0.896	1.000

Table 4 $(\omega, \alpha, \beta) = (0.3, 0.3, 0.2)$ in the log-ACD(1, 1) model

$(\omega', \alpha', \beta')$	$n = 500$	$n = 1000$	$n = 3000$
Size	0.204	0.168	0.136
(0.3, 0.3, 0.6)	1.000	1.000	1.000
(0.8, 0.3, 0.2)	1.000	1.000	1.000
(0.3, 0.1, 0.2)	0.754	0.954	1.000

Table 5 Empirical powers for hypotheses H_2 and H_3, sample size $n = 1000$

	Model	$(\omega, \alpha, \beta) \to (\omega', \alpha', \beta')$	Power
H_2	log-ACD(1, 1)	$(0.1, 0.2, 0.1) \to (0.3, 0.2, 0.1)$	0.806
H_3	log-ACD(1, 1)	$(0.1, 0.2, 0.1) \to (0.3, 0.2, 0.1)$	0.848
H_2	log-ACD(1, 1)	$(0.3, 0.3, 0.2) \to (0.3, 0.1, 0.2)$	0.910
H_3	log-ACD(1, 1)	$(0.3, 0.3, 0.2) \to (0.3, 0.1, 0.2)$	0.898

H_2: $\boldsymbol{\theta}$ changes to $\boldsymbol{\theta}'$ at $[n/3]$;
H_3: $\boldsymbol{\theta}$ changes to $\boldsymbol{\theta}'$ at $[2n/3]$.

Table 5 shows that the test produces less powers under H_2 and H_3 than under H_1.

Although we do not report here, we also experimented a simulation in the same setting as before with the nominal level 0.05. In this case, the empirical sizes appeared to be around 0.07 to 0.2 for both the ACD and log-ACD models when $n = 500$, 1000, and further, 0.06 to 0.1 when $n = 3000$. This indicates that some severe size distortions exist when the sample size is relatively small. Meanwhile, empirical powers appeared to be just a little bit less than 1.0. All these results recommend us to perform a test at the nominal level 0.1 rather than 0.05 since the former outperforms the latter. Overall, our findings in this simulation study demonstrate the validity of our cusum test.

4 Real data analysis

In this section, we apply the cusum test to analyze data from Samsung Electronics. We use observations from 9:30 to 15:30 on September 18, 2013, that is, 8045 data points with 5592 non-zero data points. Thus, we analyze 5592 non-zero data points, the mean, variance, and maximum of which are 3.219063, 9.5411644, and 43, respectively. We fit a log-ACD(1, 1) model, proposed by Bauwens and Giot (2000), to this data. To test for a change in (ω, α, β), T_n is implemented at the nominal level 0.1. To detect multiple change points, we use the ICSS algorithm (cf. Inclán and Tiao 1994). According to our analysis, four change points are found at $t = 113, 1893, 3567$, and 4631 (see the dashed line in Figs. 1, 2). The fitted models are as follows:

[1:113] $\log \psi_i = 2.216353 + 0.2265579 \log x_{i-1} - 0.5316102 \log \psi_{i-1}$, $AIC = 591.7089$;

[113:1893] $\log \psi_i = 0.1941864 + 0.08628728 \log x_{i-1} + 0.788548 \log \psi_{i-1}$, $AIC = 8278.29$;

[1893:3567] $\log \psi_i = 1.200563 - 0.005242558 \log x_{i-1} + 0.02966062 \log \psi_{i-1}$, $AIC = 7477.79$;

[3567:4631] $\log \psi_i = 1.461848 + 0.004032696 \log x_{i-1} - 0.5232957 \log \psi_{i-1}$, $AIC = 4178.785$;

[4631:5592] $\log \psi_i = 0.8072267 - 0.02545355 \log x_{i-1} + 0.02001325 \log \psi_{i-1}$, $AIC = 3479.61$.

If we ignore the change and fit the log-ACD(1, 1) model to all observations [1:5592], the fitted model is as follows:

$$\log \psi_i = 0.00734247 + 0.0280057 \log x_{i-1} + 0.972976 \log \psi_{i-1}, \quad \text{AIC} = 24048.65.$$

Comparing these AIC values also ensures the existence of change points.

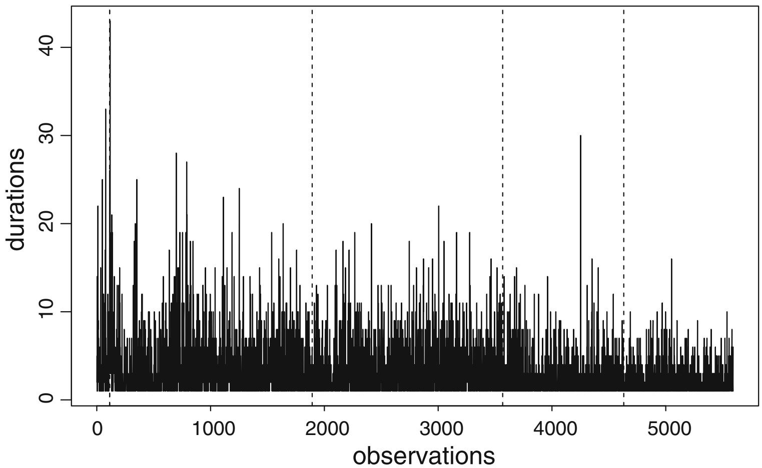

Fig. 1 Plot of the durations

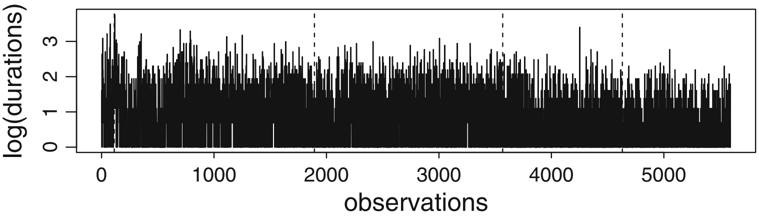

Fig. 2 Plot of the log durations

5 Concluding remarks

In this study, we investigated the cusum test based on the QMLE in nonlinear ACD models and derived its null limiting distribution under regularity conditions. To this task, we verified the strong consistency and asymptotic normality of the QMLE, and then, showed that the cusum test asymptotically behaves like the sup of the squares of independent Brownian bridges. For illustration, we implemented a simulation study and real data analysis. Na et al. (2011) studied the monitoring procedure to detect a parameter change as quickly as possible based on the cusum test. In many situations from finance to weather forecast, the online detection of parameter changes in time series models can be a crucial issue. In the same spirit, it would be of considerable interest to apply the monitoring procedure to nonlinear ACD models. Owing to its importance, we leave this as a task of our future study.

Appendix

Lemma 1 *Suppose that* (N.1)–(N.4) *hold. Then, under* H_0,

$$\mathbf{V}_0^{-\frac{1}{2}}\mathbf{C}_0^{-1}\frac{\hat{L}'_{[ns]}(\boldsymbol{\theta}_0)}{\sqrt{n}} \stackrel{w}{\longrightarrow} \mathrm{W}_d(s) \quad in \quad \mathbb{D}\left([0,1],\mathbf{R}^d\right).$$

Proof Since the random element $\psi'_i/\psi_{i,0}$ is $\mathcal{F}_{i-1}$-measurable and since $\mathcal{F}_{i-1}$ is independent of ϵ_i and $\mathbb{E}\epsilon_i = 1$, the sequence $\{l'_i(\boldsymbol{\theta}_0)\}$ forms a stationary ergodic martingale difference sequence with respect to the filtration $\{\mathcal{F}_i\}$. By (N.3), the sequence $\{l'_i(\boldsymbol{\theta}_0)\}$ is square integrable. Hence, using Theorem 18.3 of Billingsley (1999) and the Wold–Cramer device, we can express

$$\mathbf{V}_0^{-\frac{1}{2}}\mathbf{C}_0^{-1}\frac{L'_{[ns]}(\boldsymbol{\theta}_0)}{\sqrt{n}} = \mathbf{V}_0^{-\frac{1}{2}}\mathbf{C}_0^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{[ns]} l'_i(\boldsymbol{\theta}_0) \stackrel{w}{\longrightarrow} \mathrm{W}_d(s) \quad \text{in}\ \mathbb{D}\left([0,1],\mathbf{R}^d\right). \quad (6)$$

Further, (C.3) implies $\hat{\psi}_i(\boldsymbol{\theta}),\ \psi_i(\boldsymbol{\theta}) \geq \underline{g} > 0$ for all $\boldsymbol{\theta} \in K$. Hence, the mean value theorem yields

$$\|\hat{l}'_i - l'_i\|_K = \left\|\frac{\hat{\psi}'_i}{\hat{\psi}_i}\left(1-\frac{x_i}{\hat{\psi}_i}\right) - \frac{\psi'_i}{\psi_i}\left(1-\frac{x_i}{\psi_i}\right)\right\|_K$$

$$\leq C(1-x_i)\{\|\hat{\psi}_i' - \psi_i'\|_K + \|\hat{\psi}_i - \psi_i\|_K \|\psi_i'\|_K + \|\hat{\psi}_i - \psi_i\|_K \|\hat{\psi}_i' - \psi_i'\|_K\} \tag{7}$$

for some $C > 0$. Now, (7) together with an application of Lemmas 2.1 and 2.2 of Straumann and Mikosch (2006) shows that $\sum_{i=1}^{\infty} \|\hat{l}_i' - l_i'\|_K < \infty$ a.s. Hence,

$$\begin{aligned} \sup_{0\leq s\leq 1} \left\| \frac{1}{\sqrt{n}} \left(\hat{L}'_{[ns]}(\boldsymbol{\theta}_0) - L'_{[ns]}(\boldsymbol{\theta}_0) \right) \right\| &\leq \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \|\hat{l}_i'(\boldsymbol{\theta}_0) - l_i'(\boldsymbol{\theta}_0)\| \\ &\leq \frac{1}{\sqrt{n}} \sum_{i=1}^{\infty} \|\hat{l}_i' - l_i'\|_K = o(1), \quad \text{a.s.} \end{aligned} \tag{8}$$

Combining (6) and (8), we establish the lemma. □

Lemma 2 *If* (N.1)–(N.4) *hold and* $\mathbb{E}(\log^+ \|\psi_0''\|_K) < \infty$, *we have that under* H_0,

$$\frac{1}{n}\|L_n'' - \hat{L}_n''\|_K = o(1) \quad a.s.$$

Proof Similarly to (7), we can obtain

$$\begin{aligned} \|\hat{l}_i'' - l_i''\|_K &= \Bigg\| \frac{1}{\hat{\psi}_i^2} \left(\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}} \left(\frac{2x_i}{\hat{\psi}_i} - 1 \right) + \hat{\psi}_i''(\hat{\psi}_i - x_i) \right) \\ &\quad - \frac{1}{\psi_i^2} \left(\psi_i'(\psi_i')^{\mathrm{T}} \left(\frac{2x_i}{\psi_i} - 1 \right) + \psi_i''(\psi_i - x_i) \right) \Bigg\|_K \\ &\leq \left\| \frac{\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}}}{\hat{\psi}_i^2} \left(\frac{2x_i}{\hat{\psi}_i} - 1 \right) - \frac{\psi_i'(\psi_i')^{\mathrm{T}}}{\psi_i^2} \left(\frac{2x_i}{\psi_i} - 1 \right) \right\|_K \\ &\quad + \left\| \frac{\hat{\psi}_i''}{\hat{\psi}_i^2}(\hat{\psi}_i - x_i) - \frac{\psi_i''}{\psi_i^2}(\psi_i - x_i) \right\|_K \\ &\leq C(1+x_i) \Big\{ \|\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}} - \psi_i'(\psi_i')^{\mathrm{T}}\|_K + \|\hat{\psi}_i - \psi_i\|_K \|\psi_i'(\psi_i')^{\mathrm{T}}\|_K \\ &\quad + \|\hat{\psi}_i - \psi_i\|_K \|\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}} - \psi_i'(\psi_i')^{\mathrm{T}}\|_K + \|\hat{\psi}_i'' - \psi_i''\|_K \\ &\quad + \|\hat{\psi}_i - \psi_i\|_K \|\psi_i''\|_K + \|\hat{\psi}_i - \psi_i\|_K \|\hat{\psi}_i'' - \psi_i''\|_K \Big\} \end{aligned} \tag{9}$$

for some $C > 0$. Since $\mathbb{E}(\log^+ \|\psi_0'\|_K) < \infty$ implies that $\mathbb{E}(\log^+ \|\psi_0'(\psi_0')^{\mathrm{T}}\|_K) < \infty$, (9) with Lemmas 2.1 and 2.2 of the Straumann and Mikosch (2006) implies that $\|\hat{L}_n'' - L_n''\|_K \leq \sum_{i=1}^{\infty} \|\hat{l}_i' - l_i'\|_K < \infty$ a.s. This completes the proof. □

Lemma 3 *Under* (N.1)–(N.4) *with* $\mathbb{E}(\log^+ \|\psi_0''\|_K) < \infty$ *and* H_0, *we have*

$$\tilde{\mathbf{C}}_n \to \mathbf{C}_0 \quad a.s.$$

Proof Since $\mathbb{E}\|l_0''\|_K < \infty$, for any neighborhood $\mathcal{N}(\boldsymbol{\theta}_0) \subset K$, the ergodic theorem implies that

$$\frac{1}{n}\sum_{i=1}^{n} \sup_{\boldsymbol{\theta}\in\mathcal{N}(\boldsymbol{\theta}_0)} \|l_i''(\boldsymbol{\theta}) - l_i''(\boldsymbol{\theta}_0)\| \overset{\text{a.s.}}{\to} \mathbb{E} \sup_{\boldsymbol{\theta}\in\mathcal{N}(\boldsymbol{\theta}_0)} \|l_i''(\boldsymbol{\theta}) - l_i''(\boldsymbol{\theta}_0)\|.$$

Hence, for any $\epsilon > 0$, if the neighborhood $\mathcal{N}(\boldsymbol{\theta}_0)$ decreases sufficiently, we get

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \sup_{\boldsymbol{\theta}\in\mathcal{N}(\boldsymbol{\theta}_0)} \|l_i''(\boldsymbol{\theta}) - l_i''(\boldsymbol{\theta}_0)\| \le \epsilon \quad \text{a.s.}$$

Hence, for any $\epsilon > 0$, there exist a set $K'' = \{\boldsymbol{\theta} : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| \le r(\epsilon)\} \subset K$ and $N_1 \ge 1$ such that for all $n \ge N_1$,

$$\sup_{\boldsymbol{\theta}\in K''} \frac{1}{n}\|L_n''(\boldsymbol{\theta}) - L_n''(\boldsymbol{\theta}_0)\| \le \frac{1}{n}\sum_{i=1}^{n} \sup_{\boldsymbol{\theta}\in K''} \|l_i''(\boldsymbol{\theta}) - l_i''(\boldsymbol{\theta}_0)\| \le \epsilon \quad \text{a.s.}$$

Since $\hat{\boldsymbol{\theta}}_n \overset{\text{a.s.}}{\to} \boldsymbol{\theta}_0$, there exists $N_2 \ge 1$ such that $\|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\| < r(\epsilon)$ a.s. for all $n \ge N_2$. Letting $\mathbf{C}_n = -\frac{1}{n}\int_0^1 L_n''(\boldsymbol{\theta}_0 + \lambda(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0))\mathrm{d}\lambda$, we have that for $n \ge \max\{N_1, N_2\}$,

$$\begin{aligned}
\|\mathbf{C}_n - \mathbf{C}_0\| &\le \frac{1}{n}\int_0^1 \|L_n''(\boldsymbol{\theta}_0 + \lambda(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)) - L_n''(\boldsymbol{\theta}_0)\|\mathrm{d}\lambda + \left\|\frac{L_n''(\boldsymbol{\theta}_0)}{n} + \mathbf{C}_0\right\| \\
&\le \frac{1}{n}\int_0^1 \sup_{\boldsymbol{\theta}\in K''} \|L_n''(\boldsymbol{\theta}) - L_n''(\boldsymbol{\theta}_0)\|\mathrm{d}\lambda + \left\|\frac{L_n''(\boldsymbol{\theta}_0)}{n} + \mathbf{C}_0\right\| \\
&\le \epsilon + \left\|\frac{L_n''(\boldsymbol{\theta}_0)}{n} + \mathbf{C}_0\right\|.
\end{aligned}$$

This together with the fact that $L_n''(\boldsymbol{\theta}_0)/n \overset{\text{a.s.}}{\to} -\mathbf{C}_0$ implies $\|\mathbf{C}_n - \mathbf{C}_0\| = o(1)$ a.s. Further, for all $n \ge \max\{N_1, N_2\}$,

$$\begin{aligned}
\|\tilde{\mathbf{C}}_n - \mathbf{C}_n\| &\le \frac{1}{n}\int_0^1 \|L_n''(\boldsymbol{\theta}_0 + \lambda(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)) - \hat{L}_n''(\boldsymbol{\theta}_0 + \lambda(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0))\|\mathrm{d}\lambda \\
&\le \sup_{\boldsymbol{\theta}\in K''} \frac{1}{n}\|L_n''(\boldsymbol{\theta}) - \hat{L}_n''(\boldsymbol{\theta})\| \quad \text{a.s.,}
\end{aligned}$$

and thus, $\|\tilde{\mathbf{C}}_n - \mathbf{C}_n\| = o(1)$ a.s. owing to Lemma 2, which establishes the lemma. □

Lemma 4 *Suppose that* (N.1)–(N.4) *hold and* $\mathbb{E}(\log^+ \|\psi_0''\|_K) < \infty$. *Then, under* H_0,

$$\max_{1\le k\le n} \frac{\sqrt{k}}{\sqrt{n}}\|\tilde{\Delta}_k\| = o_P(1).$$

Proof Since the sequence $\{l_i'(\boldsymbol{\theta}_0)\}$ is stationary ergodic and $\mathbb{E}\|l_0'(\boldsymbol{\theta}_0)\| < \infty$, we have $\frac{1}{\sqrt{n}}\sum_{i=1}^{n^\delta}\|l_i'(\boldsymbol{\theta}_0)\| = O_p(n^{\delta-\frac{1}{2}})$ for any $\delta \in (0, \frac{1}{2})$. By the Cauchy–Schwarz inequality, we have

$$\begin{aligned}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n^\delta}\|l_i'(\boldsymbol{\theta}_0)\| - \frac{1}{\sqrt{n}}\sum_{i=1}^{n^\delta}\|\hat{l}_i'(\boldsymbol{\theta}_0)\|\right| &\le \frac{1}{\sqrt{n}}\sum_{i=1}^{n^\delta}\left|\|l_i'(\boldsymbol{\theta}_0)\| - \|\hat{l}_i'(\boldsymbol{\theta}_0)\|\right| \\ &\le \frac{1}{\sqrt{n}}\sum_{i=1}^{n^\delta}\|l_i'(\boldsymbol{\theta}_0) - \hat{l}_i'(\boldsymbol{\theta}_0)\| \\ &\le \frac{1}{\sqrt{n}}\sum_{i=1}^{\infty}\|l_i' - \hat{l}_i'\|_K = o(1) \quad \text{a.s.}\end{aligned}$$

Thus, for any $\delta \in (0, \frac{1}{2})$,

$$\max_{1\le k\le n^\delta}\left\|\frac{\hat{L}_k'(\boldsymbol{\theta}_0)}{\sqrt{n}}\right\| \le \frac{1}{\sqrt{n}}\sum_{i=1}^{n^\delta}\|\hat{l}_i'(\boldsymbol{\theta}_0)\| = o_{\mathrm{P}}(1). \tag{10}$$

Since $\tilde{\mathbf{C}}_n \stackrel{\text{a.s.}}{\to} \mathbf{C}_0$ (Lemma 3) and $\mathbf{C}_0$ is invertible, Egoroff's theorem implies that given $\epsilon > 0$ and $\eta > 0$, there exists an event A with $\mathrm{P}(A) < \epsilon$ and an $N \ge 1$, such that on A^c and for all $n \ge N$, there exists $\tilde{\mathbf{C}}_n^{-1}$ and

$$\|\tilde{\mathbf{C}}_n^{-1}\| < \|\mathbf{C}_0^{-1}\| + \eta.$$

On A^c and for all $n > N$, we have

$$\tilde{\Delta}_n = (\mathbf{C}_0 - \tilde{\mathbf{C}}_n)\tilde{\mathbf{C}}_n^{-1}\frac{\hat{L}_n'(\boldsymbol{\theta}_0)}{\sqrt{n}},$$

and further, since $\max_{k\le n^\delta}\|\mathbf{C}_0 - \tilde{\mathbf{C}}_k\| = O_p(1)$, $\max_{n^\delta\le k\le n}\|\mathbf{C}_0 - \tilde{\mathbf{C}}_k\| = o(1)$ a.s., (10) and $\max_{k\le n}\|\frac{\hat{L}_k'(\boldsymbol{\theta}_0)}{\sqrt{n}}\| \stackrel{d}{\to} \sup_{0\le s\le 1}\|\mathbf{C}_0\mathbf{V}_0^{1/2}\mathrm{W}_d(s)\|$ (Lemma 1), we have

$$\begin{aligned}\max_{N\le k\le n}\frac{\sqrt{k}}{\sqrt{n}}\|\tilde{\Delta}_k\| &= \max_{N\le k\le n}\frac{\sqrt{k}}{\sqrt{n}}\left\|(\mathbf{C}_0 - \tilde{\mathbf{C}}_k)\tilde{\mathbf{C}}_k^{-1}\frac{\hat{L}_k'(\boldsymbol{\theta}_0)}{\sqrt{k}}\right\| \\ &\le (\|\mathbf{C}_0^{-1}\| + \eta)\max_{N\le k\le n}\frac{\sqrt{k}}{\sqrt{n}}\left\|(\mathbf{C}_0 - \tilde{\mathbf{C}}_k)\frac{\hat{L}_k'(\boldsymbol{\theta}_0)}{\sqrt{k}}\right\| \\ &\le (\|\mathbf{C}_0^{-1}\| + \eta)\left\{\max_{N\le k\le n^\delta}\|(\mathbf{C}_0 - \tilde{\mathbf{C}}_k)\|\max_{N\le k\le n^\delta}\left\|\frac{\hat{L}_k'(\boldsymbol{\theta}_0)}{\sqrt{n}}\right\|\right.\end{aligned}$$

$$+ \max_{n^\delta \le k \le n} \|(\mathbf{C}_0 - \tilde{\mathbf{C}}_k)\| \max_{n^\delta \le k \le n} \left\| \frac{\hat{L}_k'(\boldsymbol{\theta}_0)}{\sqrt{n}} \right\| \Bigg\}$$
$$= o_P(1).$$

This in turn implies that for any $\epsilon > 0$,

$$\mathrm{P}\left(\max_{1 \le k \le n} \frac{\sqrt{k}}{\sqrt{n}} \|\tilde{\Delta}_k\| > \epsilon \right)$$
$$\le \mathrm{P}(A) + \mathrm{P}\left(\max_{1 \le k \le N} \frac{\sqrt{k}}{\sqrt{n}} \|\tilde{\Delta}_k\| > \epsilon \right) + \mathrm{P}\left(\max_{N \le k \le n} \frac{\sqrt{k}}{\sqrt{n}} \|\tilde{\Delta}_k\| > \epsilon,\ A^c \right).$$

This asserts the lemma. □

Lemma 5 *Under* (N.1)–(N.4) *with* $\mathbb{E}\|\psi_0'/\psi_0\|_K^2 < \infty$ *and under* H_0, $(\hat{\mathbf{V}}_0^{(n)})^{-1}$ *is a consistent estimator of* $\mathbf{V}_0^{-1}$.

Proof Since $\hat{\psi}_i(\boldsymbol{\theta}),\ \psi_i(\boldsymbol{\theta}) \ge \underline{g} > 0$ for all $\boldsymbol{\theta} \in K$, $\mathbb{E}(\log^+ \|\psi_0'\|_K) < \infty$, and $\mathbb{E}(\log^+ \|\psi_0'(\psi_0')^{\mathrm{T}}\|_K) < \infty$, using the mean value theorem and Lemmas 2.1 and 2.2 of the Straumann and Mikosch (2006), we can have

$$\left\| \frac{\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}}}{\hat{\psi}_i^2} - \frac{\psi_i'(\psi_i')^{\mathrm{T}}}{\psi_i^2} \right\|_K$$
$$\le C\left(\|\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}} - \psi_i'(\psi_i')^{\mathrm{T}}\|_K + \|\hat{\psi}_i - \psi_i\|_K \|\psi_i'(\psi_i')^{\mathrm{T}}\|_K \right) \xrightarrow{\text{e.a.s.}} 0$$

for some constant $C > 0$. Hence,

$$\left\| \frac{1}{n}\sum_{i=1}^n \frac{\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}}}{\hat{\psi}_i^2} - \frac{1}{n}\sum_{i=1}^n \frac{\psi_i'(\psi_i')^{\mathrm{T}}}{\psi_i^2} \right\|_K \le \frac{1}{n}\sum_{i=1}^\infty \left\| \frac{\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}}}{\hat{\psi}_i^2} - \frac{\psi_i'(\psi_i')^{\mathrm{T}}}{\psi_i^2} \right\|_K \to 0 \quad \text{a.s.} \tag{11}$$

Note that due to Theorem 2.7 of Straumann and Mikosch (2006) and the fact that $\mathbb{E}\|\psi_0'/\psi_0\|_K^2 < \infty$,

$$\frac{1}{n}\sum_{i=1}^n \frac{\psi_i'(\psi_i')^{\mathrm{T}}}{\psi_i^2} \xrightarrow{\text{a.s.}} \mathbb{E}\left(\frac{\psi_i'(\psi_i')^{\mathrm{T}}}{\psi_i^2} \right) \quad \text{in} \quad \mathbb{C}(K, \mathbf{R}^{d\times d}), \tag{12}$$

which together with (11) and (12) implies

$$\frac{1}{n}\sum_{i=1}^n \frac{\hat{\psi}_i'(\hat{\psi}_i')^{\mathrm{T}}}{\hat{\psi}_i^2} \xrightarrow{\text{a.s.}} \mathbb{E}\left(\frac{\psi_i'(\psi_i')^{\mathrm{T}}}{\psi_i^2} \right) \quad \text{in} \quad \mathbb{C}(K, \mathbf{R}^{d\times d}).$$

Therefore, since $\hat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$ a.s., we can conclude that

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{\psi}_i'(\hat{\boldsymbol{\theta}}_n)(\hat{\psi}_i'(\hat{\boldsymbol{\theta}}_n))^{\mathrm{T}}}{(\hat{\psi}_i(\hat{\boldsymbol{\theta}}_n))^2}\xrightarrow{\text{a.s.}}\mathbf{C}_0=\mathbb{E}\left(\frac{\psi_0'(\boldsymbol{\theta}_0)(\psi_0'(\boldsymbol{\theta}_0))^{\mathrm{T}}}{\psi_{0,0}^2}\right).$$

Similarly, it can be shown that $n^{-1}\sum_{i=1}^{n}((\hat{\epsilon}_i^{(n)})^2-1)$ converges to $\mathbb{E}(\epsilon_0^2-1)$ a.s. Hence, the lemma is validated. □

Acknowledgments We would like to thank the Editor, an AE and the three referees for their careful reading and valuable comments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012R1A2A2A01046092).

References

Bauwens, L., Giot, P. (2000). The logarithmic ACD model: An application to the bid-ask quote process of three NYSE stocks. *Annales D'Economie et de Statistique*, *60*, 117–145.
Bauwens, L., Giot, P. (2003). Asymmetric ACD models: Introducing price information in ACD models. *Empirical Economics*, *28*, 709–731.
Bauwens, L., Veredas, D. (2004). The stochastic conditional duration model: A latent variable model for the analysis of financial durations. *Journal of Econometrics*, *119*, 381–412.
Billingsley, P. (1999). *Convergence of probability measures*. New York: Wiley.
Csörgő, M., Horváth, L. (1997). *Limit theorems in change-point analysis*. New York: Wiley.
Dufour, A., Engle, R. (2000). Time and the price impact of a trade. *Journal of Finance*, *55*, 2467–2498.
Engle, R., Russell, J. (1997). Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model. *Journal of Empirical Finance*, *4*, 187–212.
Engle, R., Russell, J. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, *66*, 1127–1162.
Fernandes, M., Grammig, J. (2006). A family of autoregressive conditional duration models. *Journal of Econometrics*, *130*, 1–23.
Inclán, C., Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variances. *Journal of American Statistical Association*, *89*, 913–923.
Kang, J., Lee, S. (2014). Parameter change test for Poisson autoregressive models. *Scandinavian Journal of Statistics*, *41*, 1136–1152.
Kim, M., Lee, S. (2009). Test for tail index change in stationary time series with Pareto type marginal distribution. *Bernoulli*, *15*, 325–356.
Lee, S., Song, J. (2008). Test for parameter change in ARMA models with GARCH errors. *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*, 781–796.
Lee, S., Nishiyama, Y., Yosida, N. (2006). Test for parameter change in diffusion processes by cusum statistics based on one-step estimators. *Annals of Institute of Statistical Mathematics*, *58*, 211–222.
Meitz, M., Terasvirta, T. (2006). Evaluating models of autoregressive conditional duration. *Journal of Business and Economic Statistics*, *24*, 104–124.
Na, O., Lee, Y., Lee, S. (2011). Monitoring parameter change in time series models. *Statistical Methods and Applications*, *20*, 171–199.
Na, S., Lee, J., Lee, S. (2012). Change point test for copula ARMA–GARCH models. *Journal of Time Series Analysis*, *33*, 554–569.
Na, O., Lee, J., Lee, S. (2013). Change point detection in SCOMDY models. *AStA Advances in Statistical Analysis*, *97*, 215–238.
Page, E. (1955). A test for change in a parameter occurring at an unknown point. *Biometrika*, *42*, 523–527.
Straumann, D., Mikosch, T. (2006). Quasi maximum likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *Annals of Statistics*, *34*, 2449–2495.
Zhang, M., Russell, J., Tsay, R. (2001). A non-linear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, *104*, 179–207.