J. Japan Statist. Soc.
Vol. 34 No. 2 2004 173–188

THE CUSUM TEST FOR PARAMETER CHANGE IN REGRESSION MODELS WITH ARCH ERRORS

Sangyeol Lee*, Yasuyoshi Tokutsu** and Koichi Maekawa***

In this paper we consider the problem of testing for a parameter change in regression models with ARCH errors based on the residual cusum test. It is shown that the limiting distribution of the residual cusum test statistic is the sup of a Brownian bridge. Through a simulation study, it is demonstrated that the proposed test circumvents the drawbacks of Kim *et al.*'s (2000) cusum test. For illustration, we apply the residual cusum test to the return of yen/dollar exchange rate data.

Key words and phrases: Brownian bridge, regression models with ARCH errors, residual cusum test, test for parameter change, weak convergence.

1. Introduction

Since Page (1955), the problem of testing for a parameter change has been an important issue in statistics. It first started in the quality control context and quickly moved to other fields such as economics, engineering and medicine. So far, a large number of articles have been published in various journals. See, for instance, Brown *et al.* (1975), Wichern *et al.* (1976), Zacks (1983), Krishnaiah and Miao (1988) and Csörgő and Horváth (1997). The change point problem has drawn much attention from many researchers in time series analysis since time series often suffer from structural changes owing to changes of policy and critical social events. It is well known that detecting a change point is a crucial task and ignoring it can lead to a false conclusion. A standard example can be found in Hamilton ((1994), p. 450). For relevant references, we refer to Wichern *et al.* (1976), Picard (1985), Inclán and Tiao (1994), Mikosch and Stărică (1999), Lee and Park (2001), Lee *et al.* (2003a, b) and the papers cited in those articles.

In this paper, we concentrate ourselves on Inclán and Tiao's (1994) cusum test in regression models with ARCH errors. The ARCH and GARCH models have long been popular in financial time series analysis. For a general review, see Gouriéroux (1997). Inclán and Tiao's (1994) cusum test was originally designed for testing for variance changes and allocating their locations in iid samples. Later, it was demonstrated that the same idea can be extended to a large class of time series models (cf., Lee *et al.* (2003a)). Also, the variance change test has been studied in unstable AR models (cf., Lee *et al.* (2003b)).

In fact, Kim *et al.* (2000) considered to apply the cusum test to GARCH$(1, 1)$

Received December 2, 2003. Revised May 29, 2004. Accepted June 9, 2004.

*Department of Statistics, Seoul National University, Seoul 151-742, Korea.
**Center for Research on Regional Economic Systems, Hiroshima University, Hiroshima 730-0053, Japan.
***Faculty of Economics, Hiroshima University, Hiroshima 739-8525, Japan.

models taking account of the fact that the variance is a functional of GARCH parameters, and their change can be detected by examining the existence of the variance change. Although this reasoning was correct, it turned out that the cusum test suffers from severe size distortions and low powers. Hence, there was a demand to improve their cusum test. Here, in order to circumvent such drawbacks, we propose to use the cusum test based on the residuals, given as the squares of observations divided by estimated conditional variances. We intend to use residuals since the residual based test conventionally discard correlation effects and enhance the performance of the test. In fact, a significant improvement was observed in our simulation study.

Despite the previous work of Lee *et al.* (2003b) also considers a residual cusum test in time series models, the model of main concern was the autoregressive model with several unit roots. In fact, the mathematical analysis of the cusum test heavily relies on the probabilistic structure of the underlying time series model, and the arguments used for establishing the weak convergence result for unstable models are somewhat different from those for ARCH models. Therefore it is worth to investigate the asymptotic behavior of the residual cusum test in ARCH models. Although the present paper was originally motivated to improve Kim *et al.*'s (2000) test in the GARCH(1, 1) model, we consider the cusum test in a more general class of models including regression models with infinite order ARCH errors.

The organization of this paper is as follows. In Section 2, we introduce the residual cusum test in regression models with infinite order ARCH models that include the GARCH model, and show that its limiting distribution is the sup of a Brownian bridge. In Section 3, we perform a simulation study to compare our test with Kim *et al.*'s (2000) test in GARCH(1, 1) models. The result indicates that our method outperforms their cusum test. Then, for illustration, we apply our test to a real data set. Finally, in Section 4, we provide concluding remarks.

2. Residual cusum test

Let us consider the model

$$
\begin{aligned}
y_t &= \boldsymbol{\beta}'\boldsymbol{z}_t + \epsilon_t, \\
\epsilon_t &= h_t \xi_t, \\
h_t^2 &= a(\boldsymbol{\theta}) + \sum_{j=1}^{\infty} b_j(\boldsymbol{\theta})\epsilon_{t-j}^2,
\end{aligned}
\tag{2.1}
$$

where ξ_t are iid r.v.'s with zero mean and unit variance, $\{\boldsymbol{z}_t\}$ is a p-dimensional strictly stationary process, and $\boldsymbol{\theta} \to a(\boldsymbol{\theta})$ and $\boldsymbol{\theta} \to b(\boldsymbol{\theta})$ are nonnegative continuous real functions defined on a subset $\mathcal{N}$ in R^d with $a(\boldsymbol{\theta}) > 0$ and $\sum_{j=1}^{\infty} b_j(\boldsymbol{\theta}) < \infty$ for all $\boldsymbol{\theta} \in \mathcal{N}$. We assume that $y_s, \boldsymbol{z}_s,\ s < t$ are independent of $\xi_u, u \geq t$, and $\{(\epsilon_t, h_t, \boldsymbol{z}_t)\}$ is strong mixing. The Model (2.1) covers a broad class of important models in the financial time series context including GARCH models. In particular, it becomes a GARCH(1, 1) model if we put $\boldsymbol{z}_t = \mathbf{0}, \boldsymbol{\theta} = (\omega, \alpha_1, \alpha_2), \omega >$

$0, \alpha_1, \alpha_2 \geq 0, \alpha_1 + \alpha_2 < 1, a(\boldsymbol{\theta}) = \omega/(1 - \alpha_1 - \alpha_2)$ and $b_j(\boldsymbol{\theta}) = \alpha_1 \alpha_2^{j-1}$. In this case, $\{(\epsilon_t, h_t, \boldsymbol{z}_t)\}$ is geometrically strong mixing (cf., Carrasco and Chen (2002)). Recently, Lee and Taniguchi (2004) studied the LAN property and the residual empirical process for Model (2.1).

The objective here is to test the hypotheses

$$H_0 \ : \ \boldsymbol{\eta} = (\boldsymbol{\beta}', \boldsymbol{\theta}')' \text{ remains the same for the whole series vs.}$$
$$H_1 \ : \ \text{Not } H_0.$$

For a test, one may construct a cusum test based on $\{\hat{\varepsilon}_t := y_t - \widehat{\boldsymbol{\beta}}' \boldsymbol{z}_t\}$ as in Inclán and Tiao (1994) and Kim *et al.* (2000). However, as observed in the simulation study in Section 3, the test in GARCH(1,1) models is unstable and produces low powers. Thus one has to develop a better test which is not much affected by the GARCH parameters. As a candidate, one can naturally consider the cusum test based on $\{\xi_t^2\}$, say,

$$T_n := \frac{1}{\sqrt{n}\tau} \max_{1 \leq k \leq n} \left| \sum_{t=1}^{k} \xi_t^2 - \left(\frac{k}{n}\right) \sum_{t=1}^{n} \xi_t^2 \right|,$$

where $\tau^2 = \text{Var}(\xi_1^2)$, since T_n is free from the GARCH parameters. In this case, however, one may speculate whether T_n can detect any changes since T_n itself has no information about the GARCH parameters. But since ξ_t are not observable, one should replace ξ_t^2's by the residuals $\hat{\xi}_t^2$, which are obtained via estimating the unknown parameters. Those estimators play an important role to detect changes in the parameters in the presence of changes, while the iid property of the true errors still remains when there are no changes. From this reasoning, one can anticipate that the residual cusum test should be more stable and produce better powers.

Now, we construct the residual cusum test. To this end, we assume that

(A1) $E \parallel \boldsymbol{z}_1 \parallel^{4+\delta_1} < \infty, E|\epsilon_1|^{4+\delta_1} < \infty$ and $E|\xi_1|^{4+\delta_1} < \infty$ for some $\delta_1 > 0$.

(A2) There exists $\delta_2 > 0$ such that

$$\sup_{\|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \delta_2, \boldsymbol{\theta}' \in \mathcal{N}} \parallel \dot{a}(\boldsymbol{\theta}) \parallel < \infty \quad \text{and} \quad \sum_{j=1}^{\infty} \sup_{\|\boldsymbol{\theta} - \boldsymbol{\theta}'\| \leq \delta_2, \boldsymbol{\theta}' \in \mathcal{N}} \parallel \dot{b}_j(\boldsymbol{\theta}) \parallel < \infty,$$

where $\dot{a}(\boldsymbol{\theta})$ and $\dot{b}_j(\boldsymbol{\theta})$ denote the gradient vectors of a and b_j at $\boldsymbol{\theta}$.

(A3) There exists a sequence of positive integers with $q \to \infty, q/\sqrt{n} \to 0$ and $\sqrt{n} \sum_{j=q+1}^{\infty} b_j(\boldsymbol{\theta}) \to 0$ as $n \to \infty$.

(A4) $\{(\epsilon_t, h_t, \boldsymbol{z}_t)\}$ is strong mixing with order $\gamma(h)$ satisfying $\sum_{h=1}^{\infty} \gamma(h)^{\frac{\delta_1}{4+\delta_1}} < \infty$.

Observe that the last condition in (A3) is satisfied if $b_j(\boldsymbol{\theta})$ are geometrically bounded (as in GARCH models), and $q = [(\log n)]^{\zeta}$, $\zeta > 1$. Also, if $\boldsymbol{z}_t$ are identically zero and $\{y_t\}$ is a GARCH process, $\{(y_t, h_t)\}$ is geometrically strong mixing (cf., Carrasco and Chen (2002)), so that (A4) is satisfied.

Now, we construct the residual cusum test. In analogy of h_t^2, we define

$$\widehat{h}_t^2 = a(\widehat{\boldsymbol{\theta}}) + \sum_{j=1}^{q} b_j(\widehat{\boldsymbol{\theta}})\widehat{\epsilon}_{t-j}^2,$$
$$\widehat{\epsilon}_t = y_t - \widehat{\boldsymbol{\theta}}'\boldsymbol{z}_t \quad \text{and} \quad \widehat{\xi}_t = \widehat{\epsilon}_t/\widehat{h}_t,$$

where $\widehat{\boldsymbol{\eta}} = (\widehat{\boldsymbol{\beta}}', \widehat{\boldsymbol{\theta}}')'$ is an estimator of $\boldsymbol{\eta}$ with $\sqrt{n}(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}) = O_P(1)$. Then, we have the following result.

THEOREM 1. *Assume that* (A1)–(A4) *hold. Set*

$$\widehat{T}_n := \frac{1}{\sqrt{n}\widehat{\tau}} \max_{q+1\le k\le n} \left| \sum_{t=q+1}^{k} \widehat{\xi}_t^2 - \left(\frac{k}{n}\right) \sum_{t=q+1}^{n} \widehat{\xi}_t^2 \right|$$

where $\widehat{\tau}^2 = \frac{1}{n-q}\sum_{t=q+1}^{n}\widehat{\xi}_t^4 - (\frac{1}{n-q}\sum_{t=q+1}^{n}\widehat{\xi}_t^2)^2$. *Then, under* H_0,

$$\widehat{T}_n \xrightarrow{d} \sup_{0\le u\le 1} |B^o(u)|, \quad n \to \infty,$$

where B^o *is a Brownian bridge.*

Remark 1. A choice of q may be an issue in actual practice since it may affect the test, despite the affection would not be so serious for fairly large samples. However, if h_t^2 has a more specific form as in GARCH(1, 1) models, the test statistic can be free of the choice of q. See Theorem 2 below. In general, the above Brownian bridge result does not hold for all regression models (cf., Jandhyala and MacNeill (1991)). Therefore, the result of Theorem 1 should not be applied directly to all situations.

PROOF. Split $\widehat{\xi}_t^2$ into $\xi_t^2 + \sum_{i=1}^{6} J_{i,t}$, where

$$J_{1,t} = \frac{(h_t^2 - \widehat{h}_t^2)\xi_t^2}{h_t^2}, \qquad J_{2,t} = \frac{(h_t^2 - \widehat{h}_t^2)^2\xi_t^2}{h_t^2\widehat{h}_t^2},$$
$$J_{3,t} = \frac{-2(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{z}_t\epsilon_t}{h_t^2}, \qquad J_{4,t} = \frac{-2(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{z}_t\epsilon_t(h_t^2 - \widehat{h}_t^2)}{h_t^4},$$
$$J_{5,t} = \frac{-2(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{z}_t\epsilon_t(h_t^2 - \widehat{h}_t^2)^2}{h_t^4\widehat{h}_t^2}, \qquad J_{6,t} = \frac{((\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{z}_t)^2}{\widehat{h}_t^2}.$$

We claim that

$$\text{(2.2)} \quad \Delta_{i,n} := \frac{1}{\sqrt{n}} \max_{q+1\le k\le n} \left| \sum_{t=q+1}^{k} J_{i,t} - \left(\frac{k}{n}\right) \sum_{t=q+1}^{n} J_{i,t} \right| = o_P(1), \quad i = 1, \ldots, 6.$$

First, we handle $J_{1,t}$. Note that

$$\begin{aligned} h_t^2 - \widehat{h}_t^2 =& a(\boldsymbol{\theta}) - a(\widehat{\boldsymbol{\theta}}) + \sum_{j=q+1}^{\infty} b_j(\boldsymbol{\theta})\epsilon_{t-j}^2 \\ &+ \sum_{j=1}^{q} \left(b_j(\boldsymbol{\theta}) - b_j(\widehat{\boldsymbol{\theta}})\right)\epsilon_{t-j}^2 + \sum_{j=1}^{q} b_j(\widehat{\boldsymbol{\theta}})\left(\epsilon_{t-j}^2 - \widehat{\epsilon}_{t-j}^2\right) := \sum_{i=1}^{4} I_{i,t}. \end{aligned} \tag{2.3}$$

Owing to (A4) and the invariance principle for strong mixing processes (cf., Theorem 1.7 of Peligrad (1986)), we have

$$\frac{1}{\sqrt{n}} \max_{q+1\le k\le n} \left| \sum_{t=q+1}^{k} \left(\frac{\xi_t^2}{h_t^2} - E\frac{\xi_t^2}{h_t^2} \right) - \left(\frac{k}{n}\right) \sum_{t=q+1}^{n} \left(\frac{\xi_t^2}{h_t^2} - E\frac{\xi_t^2}{h_t^2} \right) \right| = O_P(1),$$

which implies

$$\frac{1}{\sqrt{n}} \max_{q+1\le k\le n} \left| \sum_{t=q+1}^{k} \frac{I_{1,t}\xi_t^2}{h_t^2} - \left(\frac{k}{n}\right) \sum_{t=q+1}^{n} \frac{I_{1,t}\xi_t^2}{h_t^2} \right| = o_P(1). \tag{2.4}$$

Meanwhile,

$$\frac{1}{\sqrt{n}} \max_{q+1\le k\le n} \left| \sum_{t=q+1}^{k} \frac{I_{2,t}\xi_t^2}{h_t^2} - \left(\frac{k}{n}\right) \sum_{t=q+1}^{n} \frac{I_{2,t}\xi_t^2}{h_t^2} \right| = o_P(1) \tag{2.5}$$

since by (A3),

$$\frac{1}{\sqrt{n}} \sum_{t=q+1}^{n} \sum_{j=q+1}^{\infty} b_j(\boldsymbol{\theta}) \frac{\epsilon_{t-j}^2 \xi_t^2}{h_t^2} = O_P\left(\sqrt{n} \sum_{j=q+1}^{\infty} b_j(\boldsymbol{\theta}) \right) = o_P(1).$$

Now, we verify that

$$\frac{1}{\sqrt{n}} \max_{q+1\le k\le n} \left| \sum_{t=q+1}^{k} \frac{I_{3,t}\xi_t^2}{h_t^2} - \left(\frac{k}{n}\right) \sum_{t=q+1}^{n} \frac{I_{3,t}\xi_t^2}{h_t^2} \right| = o_P(1). \tag{2.6}$$

For this task, it suffices to show that for $\lambda > 0$,

$$l_n := P\left(\sum_{j=1}^{q} |b_j(\boldsymbol{\theta}) - b_j(\widehat{\boldsymbol{\theta}})| \Lambda_{n_j} > \lambda \right) = o(1), \quad n \to \infty, \tag{2.7}$$

where

$$\Lambda_{n_j} = \frac{1}{\sqrt{n}} \max_{q+1\le k\le n} \left| \sum_{t=q+1}^{k} \left(\frac{\epsilon_{t-j}^2 \xi_t^2}{h_t^2} - E\frac{\epsilon_{t-j}^2 \xi_t^2}{h_t^2} \right) \right|$$

which is $O_P(1)$ due to the invariance principle and (A4). Observe that for any $M > 0$,

$$l_n :\leq P\left(\sum_{j=1}^{M} |b_j(\boldsymbol{\theta}) - b_j(\widehat{\boldsymbol{\theta}})|\Lambda_{n_j} > \frac{\lambda}{2}\right) + P\left(\sum_{j=M+1}^{\infty} |b_j(\boldsymbol{\theta}) - b_j(\widehat{\boldsymbol{\theta}})|\Lambda_{n_j} > \frac{\lambda}{2}\right)$$
$$:= l_{1,n} + l_{2,n},$$

$l_{1,n} = o(1)$, and

$$l_{2,n} \leq P\Bigg(\| \widehat{\boldsymbol{\theta}} - \boldsymbol{\theta} \| \times \sum_{j=M+1}^{\infty} \sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}'\|\leq\delta_2} \|\dot{b}(\boldsymbol{\theta}')\| \cdot \frac{1}{\sqrt{n}}\left(\sum_{t=1}^{n} \frac{\epsilon_{t-j}^2 \xi_t^2}{h_t^2} + \sum_{t=1}^{n} E\frac{\epsilon_{t-j}^2 \xi_t^2}{h_t^2}\right) > \frac{\lambda}{2}\Bigg)$$

for all large n. Then, using Markov's inequality and (A2), we can show that $l_{2,n}$ becomes arbitrarily small by taking a sufficiently large M. Hence, $l_{2,n} = o(1)$ and thus $l_n = o(1)$, which yields (2.6).

Now, we verify that

$$\frac{1}{\sqrt{n}} \max_{q+1\leq k\leq n} \left| \sum_{t=q+1}^{n} \frac{I_{4,t}\xi_t^2}{h_t^2} - \left(\frac{k}{n}\right) \sum_{t=q+1}^{n} \frac{I_{4,t}\xi_t^2}{h_t^2} \right| = o_P(1). \tag{2.8}$$

Note that

$$\epsilon_{t-j}^2 - \widehat{\epsilon}_{t-j}^2 = 2\epsilon_{t-j}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{z}_{t-j} - ((\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{z}_{t-j})^2.$$

Since

$$\frac{1}{\sqrt{n}} \max_{q+1\leq k\leq n} \left\| \sum_{t=q+1}^{n} \left(\frac{\boldsymbol{z}_{t-j}\epsilon_{t-j}\xi_t^2}{h_t^2} - E\frac{\boldsymbol{z}_{t-j}\epsilon_{t-j}\xi_t^2}{h_t^2} \right) \right\| = O_P(1)$$

by (A4), and

$$\begin{aligned} \sum_{j=1}^{\infty} b_j(\widehat{\boldsymbol{\theta}}) &\leq \sum_{j=1}^{\infty} \|\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\| \sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}'\|\leq\|\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\|} \left\|\dot{b}_j(\boldsymbol{\theta}')\right\| + \sum_{j=1}^{\infty} b_j(\boldsymbol{\theta}) \\ &= O_P(1), \end{aligned} \tag{2.9}$$

following essentially the same arguments between (2.6) and (2.8), we can see that

$$\begin{aligned} \frac{1}{\sqrt{n}} \max_{q+1\leq k\leq n} \Bigg\| &\sum_{t=q+1}^{k} \sum_{j=1}^{q} b_j(\widehat{\boldsymbol{\theta}})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})' \frac{\boldsymbol{z}_{t-j}\epsilon_{t-j}\xi_t^2}{h_t^2} \\ &- \left(\frac{k}{n}\right) \sum_{t=q+1}^{n} \sum_{j=1}^{q} b_j(\widehat{\boldsymbol{\theta}})(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})' \frac{\boldsymbol{z}_{t-j}\epsilon_{t-j}\xi_t^2}{h_t^2} \Bigg\| = o_P(1). \end{aligned} \tag{2.10}$$

Combining this and the fact that

$$\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n}\sum_{j=1}^{q} b_j(\widehat{\boldsymbol{\theta}})\|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\|^2\|\boldsymbol{z}_{t-j}\|^2\xi_t^2/h_t^2 = o_P(1), \quad \text{(by (2.9))}$$

we obtain (2.8). From (2.4), (2.5), (2.6) and (2.8), we establish $\Delta_{1,n} = o_P(1)$.

Now, we deal with $\Delta_{2,n}$. Since $h_t^2 \geq a(\boldsymbol{\theta}) > 0$ and $\widehat{h}_t^2 \geq a(\widehat{\boldsymbol{\theta}})$, to show $\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n} J_{2,t} = o_P(1)$, it suffices to prove

$$\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n}(h_t^2-\widehat{h}_t^2)^2\xi_t^2 = o_P(1). \tag{2.11}$$

It is obvious that $\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n} I_{1,t}^2\xi_t^2 = o_P(1)$. Also, we have

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n} I_{2,t}^2\xi_t^2 &= \frac{1}{\sqrt{n}}\sum_{t=q+1}^{n}\left(\sum_{j=q+1}^{\infty} b_j(\boldsymbol{\theta})\epsilon_{t-j}^2\right)^2\xi_t^2 \\ &= O_P\left(\sqrt{n}\left(\sum_{j=q+1}^{\infty} b_j(\boldsymbol{\theta})\right)^2\right) = o_P(1) \qquad (2.12)\end{aligned}$$

by (A3). Meanwhile, by the Cauchy-Schwarz inequality,

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n} I_{3,t}^2\xi_t^2 &\leq \frac{1}{\sqrt{n}}\sum_{t=q+1}^{n}\sum_{j=1}^{q}\|\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\|^2 \sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}'\|\leq\|\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\|}\left\|\dot{b}_j(\boldsymbol{\theta}')\right\|^2\epsilon_{t-j}^4\xi_t^4 \\ &= O_P(q/\sqrt{n}) = o_P(1). \qquad \text{(by (A3))} \qquad (2.13)\end{aligned}$$

Moreover,

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n} I_{4,t}^2\xi_t^2 &\leq \frac{2}{\sqrt{n}}\sum_{t=q+1}^{n}\left[\sum_{j=1}^{q} b_j(\boldsymbol{\theta})\{|\epsilon_{t-j}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{z}_{t-j}| + ((\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})'\boldsymbol{z}_{t-j})^2\}\right]^2\xi_t^2 \\ &= o_P(1). \qquad (2.14)\end{aligned}$$

This together with (2.11)–(2.13) yields $\Delta_{2,n} = o_P(1)$.

Now, it remains to show that $\Delta_{n,i} = o_P(1), i = 3,4,5,6$. It is trivial to show that $\Delta_{n,3} = o_P(1)$ and $\Delta_{n,6} = o_P(1)$. Also, one can verify the negligibility of $\Delta_{n,4}$ and $\Delta_{n,5}$ in a similar fashion to prove that of $\Delta_{n,1}$ and $\Delta_{n,2}$, respectively. Hence, (2.2) is established, which directly implies

$$\begin{aligned}&\frac{1}{\sqrt{n}}\max_{q+1\leq k\leq n}\left|\sum_{t=q+1}^{k}\widehat{\xi}_t^2 - \left(\frac{k}{n}\right)\sum_{t=q+1}^{n}\widehat{\xi}_t^2\right| \\ &= \frac{1}{\sqrt{n}}\max_{q+1\leq k\leq n}\left|\sum_{t=q+1}^{k}\xi_t^2 - \left(\frac{k}{n}\right)\sum_{t=q+1}^{n}\xi_t^2\right| + o_P(1). \qquad (2.15)\end{aligned}$$

Finally, we show that $\widehat{\tau}^2 \xrightarrow{P} \tau^2 = \mathrm{Var}(\xi_1^2)$. Note that

$$\widehat{\xi}_t^2 - \xi_t^2 = \frac{(h_t^2 - \hat{h}_t^2)\xi_t^2}{\widehat{h}_t^2} + \rho_t, \tag{2.16}$$

where $\rho_t := (\widehat{\varepsilon}_t^2 - \varepsilon_t^2)/\widehat{h}_t^2$ satisfies

$$\frac{1}{n}\sum_{t=q+1}^{n} \rho_t = o_P(1) \quad \text{and} \quad \frac{1}{n}\sum_{t=q+1}^{n} \rho_t^2 = o_P(1). \tag{2.17}$$

Thus, in view of (2.11) and (2.17),

$$\begin{aligned}\left|\frac{1}{n}\sum_{t=q+1}^{n}(\widehat{\xi}_t^2 - \xi_t^2)\right| &\le \left|\frac{1}{n}\sum_{t=q+1}^{n}\frac{(h_t^2 - \widehat{h}_t^2)\xi_t^2}{h_t^2}\right| + \frac{1}{n}\sum_{t=q+1}^{n}\frac{(\widehat{h}_t^2 - h_t^2)^2\xi_t^2}{h_t^2\widehat{h}_t^2} + o_P(1) \\ &\le a(\boldsymbol{\theta})\left(\frac{1}{n}\sum_{t=q+1}^{n}(h_t^2 - \widehat{h}_t^2)^2\right)^{1/2}\left(\frac{1}{n}\sum_{t=q+1}^{n}\xi_t^4\right)^{1/2} + o_P(1),\end{aligned}$$

which is $o_P(1)$ since (2.11) with ξ_t^2 replaced by 1 is also $o_P(1)$, of which proof is essentially the same as that of (2.11) and is omitted for brevity. Hence,

$$\frac{1}{n-q}\sum_{t=q+1}^{n}\widehat{\xi}_t^2 \xrightarrow{P} E\xi_1^2. \tag{2.18}$$

Now, by (2.17),

$$\begin{aligned}\frac{1}{n}\sum_{t=q+1}^{n}(\widehat{\xi}_t^2 - \xi_t^2)^2 &\le \frac{1}{n}\sum_{t=q+1}^{n}(h_t^2 - \widehat{h}_t^2)^2\xi_t^4/a(\widehat{\boldsymbol{\theta}})^2 + o_P(1) \\ &\le \left(\frac{1}{\sqrt{n}}\max_{q+1\le t\le n}\xi_t^2\right)\left(\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n}(h_t^2 - \widehat{h}_t^2)^2\xi_t^2\right)\Big/ a(\widehat{\boldsymbol{\theta}})^2 + o_P(1) \\ &= o_P(1),\end{aligned}$$

and furthermore,

$$\frac{1}{n}\sum_{t=q+1}^{n}(\widehat{\xi}_t^2 + \xi_t^2)^2 \le \frac{2}{n}\sum_{t=q+1}^{n}(\widehat{\xi}_t^2 - \xi_t^2)^2 + \frac{8}{n}\sum_{t=q+1}^{n}\xi_t^4 = O_P(1).$$

Hence,

$$\begin{aligned}\left|\frac{1}{n}\sum_{t=q+1}^{n}\widehat{\xi}_t^4 - \frac{1}{n}\sum_{t=q+1}^{n}\xi_t^4\right| &\le \left(\frac{1}{n}\sum_{t=q+1}^{n}(\widehat{\xi}_t^2 - \xi_t^2)^2\right)^{1/2}\left(\frac{1}{n}\sum_{t=q+1}^{n}(\widehat{\xi}_t^2 + \xi_t^2)^2\right)^{1/2} \\ &= o_P(1),\end{aligned}$$

so that $(n-q)^{-1}\sum_{t=q+1}^{n}\widehat{\xi}_t^4 \xrightarrow{P} E\xi_1^4$. This together with (2.18) yields $\widehat{\tau}^2 \xrightarrow{P} \tau^2$. In view of this and (2.15), we establish the theorem. □

Now, as mentioned in the remark below Theorem 1, we demonstrate that a modification of the test, free from a choice of q, can be constructed for the models with h_t^2 satisfying a specific equation. Here, considering its extreme popularity in the financial time series context, we concentrate ourselves on the case of GARCH(1, 1) errors:

$$\begin{aligned} y_t &= \boldsymbol{\beta}'\boldsymbol{z}_t + \varepsilon_t, \\ \varepsilon_t &= h_t\xi_t, \\ h_t^2 &= \omega + \alpha_1\varepsilon_{t-1}^2 + \alpha_2 h_{t-1}^2 \end{aligned} \tag{2.19}$$

with $\omega > 0, \alpha_1, \alpha_2 \geq 0$ and $\alpha_1 + \alpha_2 < 1$. In this case, we can write

$$h_t^2 = a + \alpha_1\sum_{j=1}^{\infty}\alpha_2^{j-1}\varepsilon_{t-j}^2 \tag{2.20}$$

with $a = \omega/(1-\alpha_1-\alpha_2)$, and its estimate is

$$\widehat{h}_t^2 = \widehat{a} + \widehat{\alpha}_1\sum_{j=1}^{q}\widehat{\alpha}_2^{j-1}\widehat{\varepsilon}_{t-j}^2, \tag{2.21}$$

where $\widehat{\varepsilon}_t = y_t - \widehat{\boldsymbol{\beta}}'\boldsymbol{z}_t, \widehat{\boldsymbol{\beta}}, \widehat{a}, \widehat{\alpha}_1, \widehat{\alpha}_2$ are the estimators for $\boldsymbol{\beta}, a, \alpha_1$ and α_2 satisfying

$$\begin{aligned} &\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}) = O_P(1), && \sqrt{n}(\widehat{a}-a) = O_P(1), \\ &\sqrt{n}(\widehat{\alpha}_1-\alpha_1) = O_P(1) \quad \text{and} && \sqrt{n}(\widehat{\alpha}_2-\alpha_2) = O_P(1), \end{aligned}$$

and q is a sequence of positive integers with $q \to \infty, q/\sqrt{n} \to 0$ and $\sqrt{n}\alpha_2^q \to 0$, which ensures (A3). Note that the estimate of the conditional variance can be obtained recursively from the equation

$$\tilde{h}_t^2 = \widehat{\omega} + \widehat{\alpha}_1\widehat{\varepsilon}_{t-1}^2 + \widehat{\alpha}_2\tilde{h}_{t-1}^2, \tag{2.22}$$

in so far as initial values $\widehat{\varepsilon}_0^2$ and $\tilde{h}_0^2$ are provided. From this, we have that for $t \geq 2$,

$$\tilde{h}_t^2 = \widehat{\omega}(\widehat{\alpha}_2^t - 1)/(1-\widehat{\alpha}_2) + \widehat{\alpha}_1\sum_{j=1}^{t-1}\widehat{\alpha}_2^{j-1}\widehat{\varepsilon}_{t-j}^2 + \widehat{\alpha}_1\widehat{\alpha}_2^{t-1}\widehat{\varepsilon}_0^2 + \widehat{\alpha}_2^t\tilde{h}_0^2. \tag{2.23}$$

Then, in view of (2.21) and (2.23), we have

$$\frac{1}{\sqrt{n}}\sum_{t=q+1}^{n}\widehat{\varepsilon}_t^2|\widehat{h}_t^{-2} - \tilde{h}_t^{-2}| = o_P(1), \tag{2.24}$$

and moreover,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{q}\hat{\varepsilon}_t^2|\widehat{h}_t^{-2}-\tilde{h}_t^{-2}| = O_P(q/\sqrt{n}) = o_P(1). \tag{2.25}$$

Therefore, from Theorem 1, (2.24) and (2.25), we have the following.

THEOREM 2. *Let* $\tilde{h}_t^2$ *be the one in* (2.22), *and set* $\tilde{\xi}_t^2 = \hat{\varepsilon}_t^2/\tilde{h}_t^2$. *Let*

$$\tilde{T}_n := \max_{1\le k\le n}\tilde{T}_{n,k} := \frac{1}{\sqrt{n}\tilde{\tau}}\max_{1\le k\le n}\left|\sum_{t=1}^{k}\tilde{\xi}_t^2 - \left(\frac{k}{n}\right)\sum_{t=1}^{n}\tilde{\xi}_t^2\right|,$$

where $\tilde{\tau}^2 = \frac{1}{n}\sum_{t=1}^n\tilde{\xi}_t^4 - (\frac{1}{n}\sum_{t=1}^n\tilde{\xi}_t^2)^2$. *Then if* (A1) *holds, under* H_0,

$$\tilde{T}_n \xrightarrow{d} \sup_{0\le u\le 1}|B^o(u)|, \quad n\to\infty.$$

Remark 2. Notice that unlike in $\widehat{T}_n$, the first q number of $\tilde{T}_{n,k}$'s are involved in construction of $\tilde{T}_n$. Therefore the test statistic is free from a choice of q in this sense. As for initial values $\tilde{\varepsilon}_0^2$ and $\tilde{h}_0^2$, one can put any numbers. However, one may like to choose $\tilde{\varepsilon}_0^2 = \frac{1}{n}\sum_{t=1}^n\hat{\varepsilon}_t^2$ and $\tilde{h}_0^2 = \frac{1}{n-q}\sum_{t=q+1}^n\widehat{h}_t^2$. In the latter, a choice of q is not a serious concern since initial effects somehow will disappear very fast. It may be reasoned that the initial values may affect the test, but the effect will not be severe since the last two terms in (2.23) decay to 0 exponentially fast. In the case of $\boldsymbol{z}_t = (y_{t-1},\dots,y_{t-p+1})'$, one has to adopt the test $\tilde{T}_{p,n} := \max_{p+1\le k\le n}\tilde{T}_{n,k}$ and the initial value $\tilde{\varepsilon}_{p,0}^2 = \frac{1}{n-p}\sum_{t=p+1}^n\hat{\varepsilon}_t^2$.

3. Empirical study

3.1. Simulation study

In this section, we evaluate the performance of the test statistic $\tilde{T}_n$ through a simulation study. Towards this end, we introduce the model

$$\begin{aligned} y_t &= h_t\xi_t,\\ h_t^2 &= \omega + \alpha_1 y_{t-1}^2 + \alpha_2 h_{t-1}^2, \end{aligned}$$

where y_0 is assumed to be 0 and $\{\xi_t\}$ are iid standard normal random variables. In order to see the power, we consider the following hypotheses:

$$\begin{aligned} H_0 &: \theta = (\omega,\alpha_1,\alpha_2) \text{ are constant during the time } t=1,\dots,n. \text{ vs.}\\ H_1 &: \theta \text{ changes to } \theta' = (\omega',\alpha_1',\alpha_2') \text{ at } n/2. \end{aligned}$$

Here we evaluate $\tilde{T}_n$, with $\tilde{\varepsilon}_0^2$, $\tilde{h}_0^2$ and $q = [(\log n)^2]$, for the sample size $n = 500, 800, 1000$. In particular, the $\tilde{T}_n$ is compared with Kim *et al.*'s (2000) test statistic $B_T(\hat{C})$. In this simulation we perform the test at a nominal level 0.05.

The empirical sizes and power are calculated as the rejection number of the null hypothesis out of 1000 iterations, and are summarized in Tables 1–3. The figures in the parentheses denote the sizes and powers of Kim *et al.*'s test.

As we see in the tables, our test has no severe size distortions. In particular, the test is still stable even for the case that $\alpha_1 + \alpha_2$ is close to 1 (see Tables 2 and 3). As mentioned earlier, this is because $\widehat{\xi}_t^2$ behaves asymptotically like iid ξ_t^2, unaffected by the GARCH parameters. Meanwhile, we can see that the powers are more than 0.9 at the sample size = 1000. In general, the cusum test in GARCH models requires a much larger sample size to make accurate inferences compared to iid samples. It seems that the GARCH data with volatility makes it harder to identify small changes. Compared to ours, Kim *et al.*'s test has severe size distortions and much lower powers.

Although we do not report details here, we also evaluated the test $\hat{T}_n$ with $q = [(\log n)^{3/2}], [(\log n)^2]$ and $[(\log n)^3]$. As a result, we could see that the performance of the tests with $q = [(\log n)^{3/2}]$ and $q = [(\log n)^2]$ is almost the same as the $\tilde{T}_n$, but $\hat{T}_n$ with $q = [(\log n)^3]$ performs poorly compared to the others. Actually, there is no way to choose the most optimal q. We recommend to use $[(\log n)^2]$ since it consistently gives good results in our simulation study.

Table 1. $\theta = (0.5, 0.2, 0.2)$.

$\theta' = (\omega', \alpha', \beta')$	$n = 500$	$n = 800$	$n = 1000$	$n = 1500$
Size	0.026 (0.020)	0.033 (0.025)	0.049 (0.035)	0.043 (0.039)
$(3.0, 0.2, 0.2)$	0.306 (0.077)	0.866 (0.031)	0.990 (0.009)	
$(0.5, 0.6, 0.2)$	0.493 (0.144)	0.777 (0.349)	0.901 (0.432)	
$(0.5, 0.2, 0.6)$	0.537 (0.111)	0.806 (0.269)	0.902 (0.381)	

Table 2. $\theta = (0.1, 0.4, 0.4)$.

$\theta' = (\omega', \alpha', \beta')$	$n = 500$	$n = 800$	$n = 1000$	$n = 1500$
Size	0.036 (0.009)	0.038 (0.004)	0.049 (0.005)	0.040 (0.002)
$(0.4, 0.4, 0.4)$	0.854 (0.198)	0.994 (0.387)	0.997 (0.449)	
$(0.1, 0.1, 0.4)$	0.526 (0.157)	0.839 (0.493)	0.928 (0.646)	

Table 3. $\theta = (0.1, 0.2, 0.7)$.

$\theta' = (\omega', \alpha', \beta')$	$n = 500$	$n = 800$	$n = 1000$	$n = 1500$
Size	0.020 (0.002)	0.032 (0.003)	0.032 (0.008)	0.042 (0.010)
$(0.4, 0.2, 0.7)$	0.219 (0.173)	0.722 (0.228)	0.919 (0.271)	
$(0.1, 0.2, 0.2)$	0.616 (0.070)	0.917 (0.194)	0.983 (0.313)	

Next we show an example of the simulated distribution for a estimated break point obtained by $\tilde{T}_n$, viz, the estimator of break point is the k maximizing $\tilde{T}_{n,k}$ in Theorem 2. For this task, we consider the time series that have only one structural break point in the middle of the series, i.e., $\theta = (0.1, 0.4, 0.4)$ in the first sample period is changed to $\theta' = (0.4, 0.4, 0.4)$ in the second sample period. Figures 1–3 show the distribution of estimated break points for the sample sizes

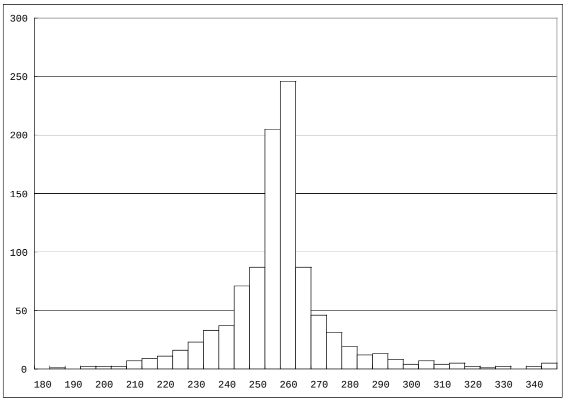

Figure 1. Estimated break point: $n = 500$.

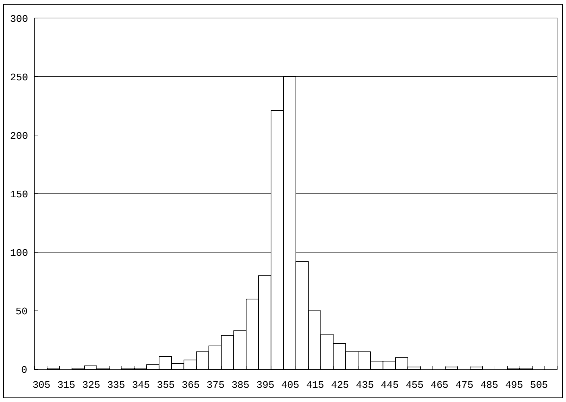

Figure 2. Estimated break point: $n = 800$.

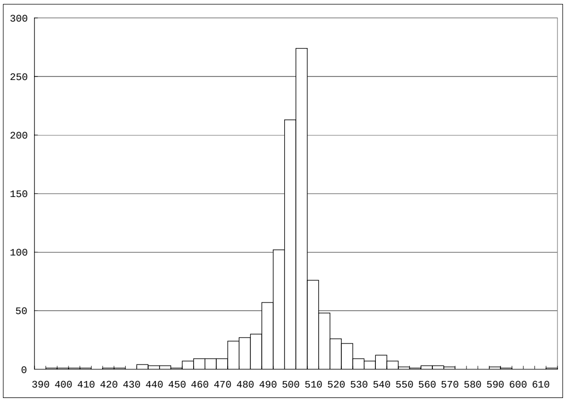

Figure 3. Estimated break point: $n = 1000$.

$n = 500, 800, 1000$, respectively. The number of iterations is 1000 for all cases. The figures indicate that the simulated distributions have a bell shape and are symmetric about the change point. The result demonstrates the validity of the estimator. Overall, our simulation study strongly supports that the residual cusum test performs adequately.

3.2. Real data analysis

In this section, we intend to demonstrate the validity of our method in actual practice. For this task, we analyze the return of yen/dollar exchange rate data from Jan. 5, 1998 to Jan. 27, 2003. Recall that the D_k plot, defined in Inclán and Tiao (1994), is a useful tool to detect multiple changes. In our case, the D_k plot is nothing but the one of $\tilde{T}_{n,k}$'s. For detecting change points, the GARCH(1,1) model is fitted to the data. Subsequently, we detected one change point on Sep.

Figure 4. Plot of Foreign Exchange rate data.

Figure 5. Plot of D_k.

28, 1999 (see the vertical line in Figures 4–5). It turns out that the data in the first period, from Jan. 5, 1998 to Sep. 28, 1999, follows the model:

$$\begin{aligned} y_t &= 0.007 + \varepsilon_t, \\ \varepsilon_t &= h_t \xi_t, \\ h_t^2 &= 0.140 + 0.175\varepsilon_{t-1}^2 + 0.686h_{t-1}^2 \end{aligned}$$

with the AIC value 1180.480, and the data in the second period follows the model

$$\begin{aligned} y_t &= 0.015 + \varepsilon_t, \\ \varepsilon_t &= h_t \xi_t, \\ h_t^2 &= 0.087 + 0.025\varepsilon_{t-1}^2 + 0.729h_{t-1}^2 \end{aligned}$$

with the AIC value 1482.389. This result indicates that the parameters experience significant changes. Unfortunately, however, we could not find any significant economic and/or political reasons for this. Meanwhile, we ignored the change on purpose and fitted the GARCH$(1,1)$ model to the whole observations. Consequently, we obtained a model very close to an IGARCH$(1,1)$ model as follows:

$$\begin{aligned} y_t &= 0.011 + \varepsilon_t, \\ \varepsilon_t &= h_t \xi_t, \\ h_t^2 &= 0.012 + 0.061\varepsilon_{t-1}^2 + 0.917h_{t-1}^2 \end{aligned}$$

with the AIC value 2686.626. The result vividly shows that ignoring changes can lead to a false conclusion in statistical inference. This misspecification result coincides with the one reported by Maekawa *et al.* (2003).

4. Concluding remarks

In this paper, we proposed a residual based cusum test based and derived that the test statistic is asymptotically distributed as the sup of a Brownian bridge under regularity conditions. In the proof, we used the invariance principle result for beta (strong) mixing processes, which was possible owing to the results of Carrasco and Chen (2002) and Peligrad (1986). The proof was of an independent interest since the mixingale approach adopted by Kim *et al.* (2000) is not easy to apply, and the proof would be much lengthier without the beta mixing condition.

In fact, the present paper was motivated to circumvent the drawbacks of the cusum test proposed by Kim *et al.* in GARCH$(1,1)$ models. The idea in developing our test is explained in Section 2. As seen in Subsection 3.1, the simulation result appeared to be remarkably favorable to our test: the sizes and powers are greatly improved compared to the original cusum test. This indicates that the residual cusum test is highly trustful. In Subsection 3.2, the test was applied to the yen/dollar exchange rate data and detected one change point. It was also seen that ignoring the change leads to a wrong conclusion. Overall, we believe that our test constitutes a functional tool for testing a parameter change in ARCH models. We leave the residual cusum test in other types of GARCH models as a topic of future study and will be reported elsewhere.

Acknowledgements

The authors thank the two referees for the valuable comments to improve the quality of the paper. The second author wishes to thank Professor Y. Suzuki for helpful discussion. The first author acknowledges that this work was supported by KOSEF in part through the Statistical Research Center for Complex System at Seoul National University. Also, the third author acknowledges the support from Grant-in-Aid for Scientific Research 14330005 by the Ministry of Education, Science and Technology, Japan.

References

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

Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models, *Econometric Theory*, **18**, 17–39.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*, John Wiley & Sons Ltd., West Sussex, England.

Giraitis, L., Kokoszka, P. and Leipus, R. (2000). Stationary ARCH models: Dependence structure and central limit theorem, *Econometric Theory*, **16**, 3–22.

Gouriéroux, C. (1997). *ARCH Model and Financial Application*, Springer, New York.

Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press, New Jersey.

Inclán, C. and Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variances, *J. Amer. Statist. Assoc.*, **89**, 913–923.

Jandhyala, V. K. and MacNeill, I. B. (1991). Tests for parameter changes at unknown times in linear regression models, *J. Statist. Plann. Infer.*, **27**, 291–316.

Kim, S., Cho, S. and Lee, S. (2000). On the cusum test for parameter changes in GARCH(1, 1) models, *Commun. Statist. Theory & Meth.*, **29**, 445–462.

Krishnaiah, P. R. and Miao, B. Q. (1988). Review about estimation of change points, *Handbook of Statistics, Vol. 7*, (eds. Krishnaiah, P. R. and Rao, C. P.), 375–402, Elsevier, New York.

Lee, S. and Park, S. (2001). The cusum of squares test for scale changes in infinite order moving average processes, *Scand. J. Statist.*, **28**, 625–644.

Lee, S. and Taniguchi, M. (2004). Asymptotic theory for ARCH-SM models: LAN and residual empirical processes, *Statistica Sinica*, (to appear).

Lee, S., Ha, J., Na, O. and Na, S. (2003a). The cusum test for parameter change in time series models, *Scand. J. Statist.*, **30**, 781–796.

Lee, S., Na, O. and Na, S. (2003b). On the cusum of squares test for variance change in nonstationary and nonparametric time series models, *Ann. Inst. Statist. Math.*, **55**, 467–485.

Maekawa, K., Lee, S. and Tokutsu, Y. (2003). Structural change and spurious volatility persistence, Discussion paper, Faculty of Economics, Hiroshima University.

Mikosch, T. and Stărică, C. (1999). Change of structure in financial time series, long range dependence and the GARCH model, Technical report, University of Groningen.

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

Peligrad, M. (1986). Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables, *Dependence in Probability and Statistics*, (eds. Eberlein, E. and Taqqu, M. S.), 193–223, Birkhäuser, Boston.

Picard, D. (1985). Testing and estimating change-points in time series, *Adv. Appl. Prob.*, **17**, 841–867.

Wichern, D. W., Miller, R. B. and Hsu, D. A. (1976). Changes of variance in first-order autoregressive time series models—with an application, *Appl. Statist.*, **25**, 248–256.

Zacks, S. (1983). Survey of classical and Bayesian approaches to the change-point problem: fixed sample and sequential procedures of testing and estimation, *Recent Advances in Statistics*, (eds. Rivzi, M. H. *et al.*), 245–269, Academic Press, New York.