Econometric Theory, **25**, 2009, 411–441. Printed in the United States of America.
doi:10.1017/S0266466608090130

ON DISTINGUISHING BETWEEN RANDOM WALK AND CHANGE IN THE MEAN ALTERNATIVES

ALEXANDER AUE
University of California, Davis

LAJOS HORVÁTH
University of Utah

MARIE HUŠKOVÁ
Charles University

SHIQING LING
Hong Kong University of Science and Technology

We study test procedures that detect structural breaks in underlying data sequences. In particular, we wish to discriminate between different reasons for these changes, such as (1) shifting means, (2) random walk behavior, and (3) constant means but innovations switching from stationary to difference stationary behavior. Almost all procedures presently available in the literature are simultaneously sensitive to all three types of alternatives.

The test statistics under investigation are based on functionals of the partial sums of observations. These cumulative sum–type (CUSUM-type) statistics have limit distributions if the mean remains constant and the errors satisfy the central limit theorem but tend to infinity in the case when any of the alternatives (1), (2), or (3) holds. On removing the effect of the shifting mean, however, divergence of the test statistics will only occur under the random walk behavior, which in turn enables statisticians not only to detect structural breaks but also to specify their causes.
The results are underlined by a simulation study and an application to returns of the German stock index DAX.

1. INTRODUCTION

Since the contribution of Nelson and Plosser (1982), developing tests to distinguish between the stationarity of the errors and the stationarity of the error

We are grateful to M.L. Reimherr for providing the simulations and P. Kokoszka for making available the data used in Section 4. This research was partially supported by NSF grants DMS 0604670 and DMS 0652420, grant RGC-HKUST6428/06H, grant MSM 0021620839, and grant GAČR 201/06/0186. Address correspondence to Alexander Aue, Department of Statistics, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA; e-mail: alexaue@wald.ucdavis.edu.

© 2009 Cambridge University Press 0266-4666/09 $15.00

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

increments (difference stationarity) has become an integral part of econometrics. The list of publications in the area includes, among many others, Perron (1989, 1990) and Perron and Vogelsang (1992), who considered unit-root tests in a mean change environment, and Andreau and Spanos (2003), who revisited the empirical results in Nelson and Plosser (1982) and Perron (1989) and reported that some models introduced in the latter papers were misspecified. Instead, Andreau and Spanos (2003) argued that the majority of the data series are trend stationary.

Recently, there has also been increased interest in testing for changes in the mean versus difference stationarity and testing for changes from stationarity to difference stationarity. Nyblom and Mäkeläinen (1983) and Kwiatkowski, Phillips, Schmidt, and Shin (1992) proposed tests for stationarity based on partial sums of the observations. Related tests were introduced by Banerjee, Lumsdaine, and Stock (1992), Leybourne, Kim, Smith, and Newbold (2003), Xiao (2001), and Harvey and Mills (2004). All procedures, however, were designed to detect difference stationary errors and have no power against changes in the mean and strongly dependent observations (see Giraitis, Kokoszka, and Leipus 2001; Giraitis, Kokoszka, Leipus, and Teyssiere, 2003). A method to distinguish between long-range dependence and changes in the mean was provided by Berkes, Horváth, Kokoszka, and Shao (2006). Kim (2000) (see also Kim, Belaire-Franch, and Badillo-Amado, 2002) proposed ratio-based procedures to test stationarity against changing persistence. Busetti and Taylor (2004) and Taylor (2005) considered modifications of ratio-based tests and their properties under different sets of assumptions. Belaire-Franch (2005) proved that the Kim (2000) test is unable to differentiate between shifting means and changes from stationarity to difference stationarity.

In this paper we investigate procedures that test for the presence of a structural break in the underlying data and extract whether this break is caused by (1) a shifting mean, (2) random walk behavior, or (3) errors switching from stationary to difference stationary behavior under constant means. To do so, we use functionals of partial sums of the observations to construct these tests, which are based on versions of the cumulative sum (CUSUM) statistics (see Page, 1954), the rescaled range (R/S) statistics (see Lo, 1991), and the rescaled variance (V/S) statistic (see Giraitis et al., 2001). These have limit distributions if the mean is constant and the sums of the errors satisfy the central limit theorem. Under the alternatives (1)–(3), all test statistics tend to infinity in probability. After removing the effect of a possibly shifting mean, however, the procedures will diverge only under the random walk–type alternatives (2) and (3), so that discrimination between these different phenomena is achieved.

The paper is organized as follows. In Section 2, we carefully introduce the model by describing in some detail the different alternatives under consideration, defining the test statistics, and stating the main results. Section 3 contains a simulation study that reports the finite-sample properties of the testing procedures, and Section 4 addresses an application to the daily returns of the German stock index DAX. The mathematical proofs of all main results are provided in Section 5.

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

2. RESULTS

We study a sequence of random variables $\{X_k\}$ given by

$$X_k = \mu_k + Y_k, \qquad k = 1, \ldots, n,$$

under four distinct scenarios, which require different assumptions on the trend parameters $\{\mu_k\}$ and the error terms $\{Y_k\}$. Scenario I describes the case of structural stability along with short memory innovations, Scenario II presents a typical change-point model with (at most one) shift in the mean but short memory innovations, whereas for Scenario III (Scenario IV) the trend is constant with random walk–type errors (with errors changing into a random walk). Scenario I will be used as null hypothesis; the other three serve as alternatives. The precise formulations are as follows.

SCENARIO I. Under the null hypothesis of structural stability, the trend terms $\{\mu_k\}$ remain constant, i.e.,

$$\mathcal{H}_0 : \ \mu_k = \mu \quad \text{for all } k = 1, \ldots, n. \tag{2.1}$$

It is also assumed that the error terms $\{Y_k\}$ satisfy the functional central limit theorem (FCLT), i.e., with some $\sigma > 0$,

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} Y_k \overset{\mathcal{D}[0,1]}{\to} \sigma W(t) \qquad (n \to \infty), \tag{2.2}$$

where $\lfloor \cdot \rfloor$ is the integer part, $\{W(t) : t \in [0, 1]\}$ denotes a standard Brownian motion (Wiener process), and $\overset{\mathcal{D}[0,1]}{\to}$ stands for weak convergence in the Skorohod space $\mathcal{D}[0, 1]$.

SCENARIO II. Under the first alternative, the mean shifts from μ to $\mu + \Delta$ at an unknown time k^*, the so-called change-point, whereas the innovation sequence $\{Y_k\}$ remains weakly dependent, i.e.,

$$\mathcal{H}_A^{(1)} : \quad \mu_k = \begin{cases} \mu, & k = 1, \ldots, k^*, \\ \mu + \Delta, & k = k^* + 1, \ldots, n, \end{cases} \quad \text{and } \{Y_k\} \text{ satisfies (2.2).} \tag{2.3}$$

Associated with $\mathcal{H}_A^{(1)}$ is a broad body of literature, underlining the importance of the topic. For a comprehensive review on the subject refer to Csörgő and Horváth (1997). The parameter Δ is usually unknown, possibly depending on n.

SCENARIO III. Under the second alternative, the mean remains constant, but the error terms themselves (and not their partial sums) satisfy the FCLT, i.e.,

$$\mathcal{H}_A^{(2)} : \quad \mu_k = \mu \quad \text{for all } k = 1, \ldots, n, \tag{2.4}$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

$$\frac{1}{\sqrt{n}} Y_{\lfloor nt \rfloor} \overset{\mathcal{D}[0,1]}{\to} \sigma W(t) \qquad (n \to \infty) \tag{2.5}$$

with some $\sigma > 0$ and a Brownian motion $\{W(t) : t \in [0,1]\}$. Here, $\{Y_k\}$ behaves like a random walk.

SCENARIO IV. Under the third alternative, the mean remains stable, and at time k^* the stationary Y_k changes into a random walk, i.e.,

$$\mathcal{H}_A^{(3)}: \quad \mu_k = \mu \quad \text{for all } k = 1, \ldots, n, \tag{2.6}$$

$$\text{the partial sums of } Y_1, \ldots Y_{k^*} \text{ satisfy the FCLT,} \tag{2.7}$$

$$\text{and} \quad Y_{k^*+1}, \ldots, Y_n \text{ satisfy (2.5).} \tag{2.8}$$

In the case of Scenario IV the asymptotically stationary innovations turn into a random walk at the unknown time k^*.

Without imposing overly restrictive assumptions, the limit relations (2.2) and (2.5) are satisfied by almost all important classes of innovation sequences important in applications, such as martingale differences (see Hall and Heyde, 1980), mixing sequences (see Philipp, 1986; Shao, 1993), and processes displaying heteroskedasticity (see Carrasco and Chen, 2002; Aue, Berkes, and Horváth, 2006).

It is well known that most of the test statistics used to detect changes in the mean are sensitive to other phenomena such as random walk behavior. In addition, test procedures designed to discover random walks (e.g., unit-root tests) also reject in the presence of mean changes. In what follows, we introduce two-step procedures—based on different test statistics—that are able to discriminate between the scenarios under consideration. Before doing so, we determine an estimator for the unknown asymptotic variance of the partial sums constructed from the sequence $\{X_k\}$. Because of the possible dependence between the observations, we use the Bartlett estimator defined by

$$s_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 + \frac{2}{n} \sum_{j=1}^{q(n)} w_j(q(n)) \sum_{i=1}^{n-j} (X_i - \bar{X}_n)(X_{i+j} - \bar{X}_n),$$

where the Bartlett weights w_j are given by

$$w_j(q(n)) = 1 - \frac{j}{q(n)+1}$$

with an increasing sequence $\{q(n)\}$ for which the conditions

$$q(0) > 0 \quad \text{and} \quad q(n) \to \infty, \qquad \frac{q(n) \log n}{n} \to 0 \qquad (n \to \infty) \tag{2.9}$$

are satisfied. The almost sure asymptotics of s_n^2, which is going to be used in the present paper, has been established in Berkes, Horváth, Kokoszka, and Shao

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

(2005). Note that, in the case of weakly dependent processes, s_n^2 estimates the variance of the sample mean (or the spectral density at zero frequency; see, e.g., Bartlett, 1950; Andrews, 1991).

For $n \geq 1$, let

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

denote the sample mean of the first n observations. We are interested in a CUSUM-type procedure and test statistics based on rescaled range and variance analysis, respectively. The definitions are as follows.

PROCEDURE I. The popular CUSUM statistic is given by

$$T_n = \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} (X_i - \bar{X}_n) \right|.$$

Observe that T_n compares the differences in the sample means up to time k with the sample mean obtained from the whole observation period of length n. Procedures based on T_n are intensively used in change-point analysis and were introduced originally in the field of quality control (see Page, 1954).

PROCEDURE II. The so-called rescaled range analysis, originally aimed at detecting long-range dependence, was introduced by Hurst (1951) and was subsequently refined by several authors (see Mandelbrot and Taqqu, 1979, and references therein). Here, we are interested in a modified version of the R/S statistic proposed by Lo (1991), which is based on the adjusted range

$$R_n = \frac{1}{\sqrt{n}} \left(\max_{1 \leq k \leq n} \sum_{i=1}^{k} (X_i - \bar{X}_n) - \min_{1 \leq k \leq n} \sum_{i=1}^{k} (X_i - \bar{X}_n) \right).$$

There are numerous contributions examining long memory in speculative assets such as stock returns using rescaled range analysis. We only cite Teverovsky, Taqqu, and Willinger (1999) and Breidt, Crato, and de Lima (1998) and refer to the references in these papers.

PROCEDURE III. A modification of the R/S statistic was suggested by Giraitis et al. (2001). It is defined via the comparison of sample variances rather than the adjusted ranges. Here,

$$V_n = \frac{1}{n^2} \left(\sum_{k=1}^{n} \left[\sum_{j=1}^{k} (X_j - \bar{X}_n) \right]^2 - \frac{1}{n} \left[\sum_{k=1}^{n} \sum_{j=1}^{k} (X_j - \bar{X}_n) \right]^2 \right).$$

Note that this so-called V/S statistic can be obtained from the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test introduced by Kwiatkowski et al. (1992)

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

by a mean correction. It turned out that this modification yields somewhat nicer asymptotic and finite-sample properties (see Giraitis et al., 2001, 2003).

First, we consider the asymptotics of all test procedures under the null hypothesis $\mathcal{H}_0$. As an additional condition we assume the convergence in probability of the Bartlett estimator s_n that is used to normalize T_n, R_n, and V_n, because conditions (2.1), (2.2), and (2.9) alone may not be strong enough to imply (2.11). We refer to Theorem 3.1 in Giraitis et al. (2003) for a list of sufficient criteria. The second part of this theorem specifies the behavior of the change-point estimator

$$\hat{k}(n) = \min\left\{k = 1, \ldots, n : T_n = \frac{1}{\sqrt{m}}\left|\sum_{i=1}^{k}(X_i - \bar{X}_n)\right|\right\} \tag{2.10}$$

under $\mathcal{H}_0$. If there is at most one change of the form $k^* = \lfloor \theta n \rfloor$ for some $\theta \in (0, 1)$, then $\hat{\theta}(n) = \hat{k}(n)/n$ provides a consistent and asymptotically unbiased estimator for θ under the conditions imposed under Scenario II. See Csörgő and Horváth (1997) for details.

THEOREM 2.1 (Asymptotics under $\mathcal{H}_0$). *If conditions (2.1) and (2.2) are satisfied, and if*

$$s_n^2 \xrightarrow{P} \sigma^2 \qquad (n \to \infty), \tag{2.11}$$

then, as $n \to \infty$,

$$\frac{T_n}{s_n} \xrightarrow{\mathcal{D}} \sup_{0 \le t \le 1} |B(t)|, \tag{2.12}$$

$$\frac{R_n}{s_n} \xrightarrow{\mathcal{D}} \sup_{0 \le t \le 1} B(t) - \inf_{0 \le t \le 1} B(t), \tag{2.13}$$

$$\frac{V_n}{s_n^2} \xrightarrow{\mathcal{D}} \int_0^1 B^2(t)\,dt - \left(\int_0^1 B(t)\,dt\right)^2, \tag{2.14}$$

where $\{B(t) : t \in [0, 1]\}$ *denotes a Brownian bridge. Also, as* $n \to \infty$,

$$\frac{\hat{k}(n)}{n} \xrightarrow{P} \arg\max_{t \in [0,1]} |B(t)| = \min\left\{s \in [0, 1] : |B(s)| = \sup_{t \in [0,1]} |B(t)|\right\}. \tag{2.15}$$

The limit in (2.12) is equal to the limit of the Kolmogorov–Smirnov statistic. In (2.13), we have the range of a Brownian bridge (see Feller, 1951), which is also the limit of the Kuiper statistic (see Kuiper, 1960). The distribution function of the limit in (2.14) was obtained by Watson (1961). Critical values of all three limit distributions are provided on pages 143, 144, and 148 of Shorack and Wellner (1986).

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

Next, we study the behavior of the test statistics under the change in the mean alternative $\mathcal{H}_A^{(1)}$. We assume that the change-point has the specific form

$$k^* = \lfloor \theta n \rfloor \quad \text{with some } \theta \in (0,1) \tag{2.16}$$

and that the conditions

$$n\Delta^2 \to \infty \quad \text{and} \quad \Delta^2 |\hat{k}(n) - k^*| = \mathcal{O}_P(1) \qquad (n \to \infty) \tag{2.17}$$

are satisfied and $\Delta = \Delta_n$ is a function of the sample size n. Condition (2.16) is a typical assumption on the change-point that is needed for limit theorems, such as weak and strong approximations, to apply. It assures, roughly, that the number of observations before and after the change both tend to infinity with increasing sample size. Hence, early and late changes are excluded from the discussions to come.

The second part of (2.17) is in particular fulfilled if the observations are uncorrelated moving averages (see Bai, 1994), strictly stationary sequences satisfying the FCLT (2.2) (see Csörgő and Horváth, 1997, Thm. 4.1.4), or squares of ARCH(∞) processes (see Giraitis et al., 2001).

We arrive at the following theorem.

THEOREM 2.2 (Asymptotics under $\mathcal{H}_A^{(1)}$). *If conditions (2.2), (2.3), (2.16), and (2.17) are satisfied, and if*

$$\Delta^2 q(n) = \mathcal{O}(1) \qquad (n \to \infty) \tag{2.18}$$

holds, then, as $n \to \infty$,

$$\frac{T_n}{s_n} \xrightarrow{P} \infty, \tag{2.19}$$

$$\frac{R_n}{s_n} \xrightarrow{P} \infty, \tag{2.20}$$

$$\frac{V_n}{s_n^2} \xrightarrow{P} \infty. \tag{2.21}$$

Under the alternative $\mathcal{H}_A^{(2)}$, i.e., if the mean remains unchanged but the innovations change to random walk behavior, we obtain the following limiting behavior.

THEOREM 2.3 (Asymptotics under $\mathcal{H}_A^{(2)}$). *If conditions (2.4), (2.5), and (2.9) are satisfied, then, as* $n \to \infty$,

$$\sqrt{\frac{q(n)}{n}} \frac{T_n}{s_n} \xrightarrow{\mathcal{D}} \left(\int_0^1 \Psi^2(x)\,dx \right)^{-1/2} \sup_{0 \le t \le 1} \left| \int_0^t \Psi(x)\,dx \right|,$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

$$\sqrt{\frac{q(n)}{n}}\frac{R_n}{s_n} \xrightarrow{\mathcal{D}} \left(\int_0^1 \Psi^2(x)\,dx\right)^{-1/2}\left(\sup_{0\le t\le 1}\int_0^t \Psi(x)\,dx - \inf_{0\le t\le 1}\int_0^t \Psi(x)\,dx\right),$$

$$\frac{q(n)}{n}\frac{V_n}{s_n^2} \xrightarrow{\mathcal{D}} \left(\int_0^1 \Psi^2(x)\,dx\right)^{-1} \times \left(\int_0^1 \left[\int_0^t \Psi(x)\,dx\right]^2 dt - \left[\int_0^1\int_0^t \Psi(x)\,dx\,dt\right]^2\right),$$

where

$$\Psi(x) = W(x) - \int_0^1 W(u)du, \qquad x \in [0,1].$$

Next, we consider the case when the asymptotically stationary innovations turn into a random walk.

THEOREM 2.4 (Asymptotics under $\mathcal{H}_A^{(3)}$). *If conditions (2.4) and (2.9) are satisfied,*

$$\frac{k^*}{n} \to \theta \qquad (n\to\infty) \quad \textit{and} \quad 0 \le \theta < 1, \tag{2.22}$$

then, as $n \to \infty$,

$$\sqrt{\frac{q(n)}{n}}\frac{T_n}{s_n} \xrightarrow{\mathcal{D}} \left(\int_0^1 \Psi_\theta^2(x)\,dx\right)^{-1/2}\sup_{0\le t\le 1}\left|\int_0^t \Psi_\theta(x)\,dx\right|,$$

$$\sqrt{\frac{q(n)}{n}}\frac{R_n}{s_n} \xrightarrow{\mathcal{D}} \left(\int_0^1 \Psi_\theta^2(x)\,dx\right)^{-1/2}\left(\sup_{0\le t\le 1}\int_0^t \Psi_\theta(x)\,dx - \inf_{0\le t\le 1}\int_0^t \Psi_\theta(x)\,dx\right),$$

$$\frac{q(n)}{n}\frac{V_n}{s_n^2} \xrightarrow{\mathcal{D}} \left(\int_0^1 \Psi_\theta^2(x)\,dx\right)^{-1}\left(\int_0^1 \left[\int_0^t \Psi_\theta(x)\,dx\right]^2 dt - \left[\int_0^1\int_0^t \Psi_\theta(x)\,dx\,dt\right]^2\right),$$

where

$$\Psi_\theta(x) = \begin{cases} -\displaystyle\int_\theta^1 W(u-\theta)\,du, & x \in [0,\theta] \\ W(x-\theta) - \displaystyle\int_\theta^1 W(u-\theta)\,du, & x \in [\theta,1]. \end{cases}$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

Condition (2.22) means that a fraction $1-\theta$ of the data is a random walk.

Note that, under Theorems 2.2–2.4, all test statistics under consideration satisfy

$$\frac{T_n}{s_n} \xrightarrow{P} \infty, \qquad \frac{R_n}{s_n} \xrightarrow{P} \infty, \quad \text{and} \quad \frac{V_n}{s_n^2} \xrightarrow{P} \infty \qquad (n \to \infty).$$

Consequently, the CUSUM, rescaled range, and rescaled variance tests are sensitive under both alternatives. However, neither of these procedures can, in its present form, discriminate between changes in the mean along with stationary errors and a constant mean along with random walk–type errors.

But, if the mean change is taken into account in the construction of the testing procedures, under $\mathcal{H}_A^{(1)}$ the data will be stationary. Modified test statistics will be defined utilizing the change-point estimator $\hat{k} = \hat{k}(n)$ given in (2.10). The basic idea is to split the data into two samples at $\hat{k}$, recognizing that the first sample $X_1, \ldots, X_{\hat{k}}$ and the second sample $X_{\hat{k}+1}, \ldots, X_n$ follow the same pattern under $\mathcal{H}_A^{(1)}$. To this end, let

$$T_{n,1} = \max_{1 \le k \le \hat{k}} \frac{1}{\sqrt{\hat{k}}} \left| \sum_{i=1}^{k} X_i - \frac{k}{\hat{k}} \sum_{i=1}^{\hat{k}} X_i \right|,$$

$$T_{n,2} = \max_{\hat{k} < k \le n} \frac{1}{\sqrt{n-\hat{k}}} \left| \sum_{i=\hat{k}+1}^{k} X_i - \frac{k-\hat{k}}{n-\hat{k}} \sum_{i=\hat{k}+1}^{n} X_i \right|$$

be the CUSUM procedures, respectively, before and after the estimated change. The new statistic is obtained by determining the maximum of (normalized versions of) $T_{n,1}$ and $T_{n,2}$, i.e.,

$$T_n^* = \max \left\{ \frac{T_{n,1}}{s_{n,1}}, \frac{T_{n,2}}{s_{n,2}} \right\}, \tag{2.23}$$

where $s_{n,1}^2$ and $s_{n,2}^2$ are the corresponding Bartlett estimators based on the first $\hat{k}$ observations and the last $n-\hat{k}$ observations, respectively. That is,

$$s_{n,1}^2 = \frac{1}{\hat{k}} \sum_{i=1}^{\hat{k}} (X_i - \bar{X}_{\hat{k}})^2 + 2 \sum_{i=1}^{q(\hat{k})} w_i(q(\hat{k})) \frac{1}{\hat{k}} \sum_{j=1}^{\hat{k}-i} (X_j - \bar{X}_{\hat{k}})(X_{i+j} - \bar{X}_{\hat{k}}),$$

$$s_{n,2}^2 = \frac{1}{n-\hat{k}} \sum_{i=\hat{k}+1}^{n} (X_i - \tilde{X}_{\hat{k}})^2$$

$$+ 2 \sum_{i=1}^{q(n-\hat{k})} w_i(q(n-\hat{k})) \frac{1}{n-\hat{k}} \sum_{j=\hat{k}+1}^{n-i} (X_j - \tilde{X}_{\hat{k}})(X_{i+j} - \tilde{X}_{\hat{k}})$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

with

$$\tilde{X}_\ell = \frac{1}{n-\ell} \sum_{i=\ell+1}^{n} X_i, \qquad \ell = 1, \ldots, n-1,$$

being the sample mean of the last $n-\ell$ observations. Similarly, we define the modified versions of the R/S statistic R_n and the V/S statistic V_n by setting

$$R_{n,1} = \frac{1}{\sqrt{\hat{k}}} \left(\max_{1 \le k \le \hat{k}} \sum_{i=1}^{k} (X_i - \bar{X}_{\hat{k}}) - \min_{1 \le k \le \hat{k}} \sum_{i=1}^{k} (X_i - \bar{X}_{\hat{k}}) \right),$$

$$R_{n,2} = \frac{1}{\sqrt{n-\hat{k}}} \left(\max_{\hat{k} < k \le n} \sum_{i=\hat{k}+1}^{k} (X_i - \tilde{X}_{\hat{k}}) - \min_{\hat{k} < k \le n} \sum_{i=\hat{k}+1}^{k} (X_i - \tilde{X}_{\hat{k}}) \right)$$

with

$$R_n^* = \max \left\{ \frac{R_{n,1}}{s_{n,1}}, \frac{R_{n,2}}{s_{n,2}} \right\} \tag{2.24}$$

and

$$V_{n,1} = \frac{1}{\hat{k}^2} \left(\sum_{k=1}^{\hat{k}} \left[\sum_{j=1}^{k} (X_j - \bar{X}_{\hat{k}}) \right]^2 - \frac{1}{\hat{k}} \left[\sum_{k=1}^{\hat{k}} \sum_{j=1}^{k} (X_j - \bar{X}_{\hat{k}}) \right]^2 \right),$$

$$V_{n,2} = \frac{1}{(n-\hat{k})^2} \left(\sum_{k=\hat{k}+1}^{n-\hat{k}} \left[\sum_{j=1}^{k} (X_j - \tilde{X}_{\hat{k}}) \right]^2 - \frac{1}{n-\hat{k}} \left[\sum_{k=\hat{k}+1}^{n} \sum_{j=1}^{k} (X_j - \tilde{X}_{\hat{k}}) \right]^2 \right)$$

with

$$V_n^* = \max \left\{ \frac{V_{n,1}}{s_{n,1}^2}, \frac{V_{n,2}}{s_{n,2}^2} \right\}. \tag{2.25}$$

These newly designed statistics are able to distinguish between our two alternatives $\mathcal{H}_A^{(1)}$ and $\mathcal{H}_A^{(2)}$ (similarly, between $\mathcal{H}_A^{(1)}$ and $\mathcal{H}_A^{(3)}$), which can be seen from the next set of results.

THEOREM 2.5 (Asymptotics under $\mathcal{H}_A^{(1)}$). *If conditions (2.3), (2.16), and (2.17) are satisfied, then, as* $\to \infty$,

$$(T_{n,1}, T_{n,2}) \xrightarrow{\mathcal{D}} \sigma \left(\sup_{0 \le t \le 1} |B^{(1)}(t)|, \sup_{0 \le t \le 1} |B^{(2)}(t)| \right),$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

$$(R_{n,1}, R_{n,2}) \xrightarrow{\mathcal{D}} \sigma\left(\sup_{0\le t\le 1} B^{(1)}(t) - \inf_{0\le t\le 1} B^{(1)}(t),\ \sup_{0\le t\le 1} B^{(2)}(t) - \inf_{0\le t\le 1} B^{(2)}(t)\right),$$

$$(V_{n,1}, V_{n,2}) \xrightarrow{\mathcal{D}} \sigma^2\left(\int_0^1 (B^{(1)})^2(t)\,dt - \left[\int_0^1 B^{(1)}(t)\,dt\right]^2,\right.$$

$$\left.\int_0^1 (B^{(2)})^2(t)\,dt - \left[\int_0^1 B^{(2)}(t)\,dt\right]^2\right),$$

where $\{B^{(1)}(t), t \in [0,1]\}$ *and* $\{B^{(2)}(t), t \in [0,1]\}$ *denote independent Brownian bridges.*

As an immediate consequence of the continuous mapping theorem, Theorem 2.5 implies the convergence of T_n^*, R_n^*, and V_n^* under $\mathcal{H}_A^{(1)}$, provided the Bartlett estimators $s_{n,1}^2$ and $s_{n,2}^2$ converge.

COROLLARY 2.1 (Asymptotics under $\mathcal{H}_A^{(1)}$). *If condition (2.3) is satisfied,*

$$s_{n,1}^2 \xrightarrow{P} \sigma^2 \quad \textit{and} \quad s_{n,2}^2 \xrightarrow{P} \sigma^2, \tag{2.26}$$

then, as $n \to \infty$,

$$T_n^* \xrightarrow{\mathcal{D}} \max\left\{\sup_{0\le t\le 1} |B^{(1)}(t)|,\ \sup_{0\le t\le 1} |B^{(2)}(t)|\right\},$$

$$R_n^* \xrightarrow{\mathcal{D}} \max\left\{\sup_{0\le t\le 1} B^{(1)}(t) - \inf_{0\le t\le 1} B^{(1)}(t),\ \sup_{0\le t\le 1} B^{(2)}(t) - \inf_{0\le t\le 1} B^{(2)}(t)\right\},$$

$$V_n^* \xrightarrow{\mathcal{D}} \max\left\{\int_0^1 (B^{(1)})^2(t)\,dt - \left[\int_0^1 B^{(1)}(t)\,dt\right]^2,\right.$$

$$\left.\int_0^1 (B^{(2)})^2(t)\,dt - \left[\int_0^1 B^{(2)}(t)\,dt\right]^2\right\},$$

where $\{B^{(1)}(t), t \in [0,1]\}$ *and* $\{B^{(2)}(t), t \in [0,1]\}$ *denote independent Brownian bridges.*

According to the discussion after Theorem 2.1, the distribution functions in the limits in Corollary 2.1 can be computed very easily because they are the maxima of two independent, identically distributed random variables with known distribution functions (see Shorack and Wellner, 1986, pp. 143, 144, 148).

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

To establish (2.26) delicate arguments are needed such as proving the convergence in probability of a randomly stopped sum. Berkes et al. (2005, 2006) have provided sufficient conditions.

Finally, we show that all three modified statistics are asymptotically consistent against the random walk alternative $\mathcal{H}_A^{(2)}$ and changing into a random walk alternative $\mathcal{H}_A^{(3)}$.

THEOREM 2.6 (Asymptotics under $\mathcal{H}_A^{(2)}$). *If conditions (2.4), (2.5), and (2.9) are satisfied, then, as* $n \to \infty$,

$$T_n^* \xrightarrow{P} \infty, \qquad R_n^* \xrightarrow{P} \infty, \quad \textit{and} \quad V_n^* \xrightarrow{P} \infty. \tag{2.27}$$

THEOREM 2.7 (Asymptotics under $\mathcal{H}_A^{(3)}$). *If conditions (2.6)–(2.9) are satisfied, then (2.27) holds.*

It is immediate from the proofs that the procedures are consistent against any alternatives for which a fraction $0 < \beta \leq 1$ of the data is a random walk.

Finally, we mention briefly possible extensions to scenarios allowing for more than one change-point. Given that the number of changes can be controlled by an upper bound (see Yao, 1988), the so-called binary segmentation method developed in Vostrikova (1981) applies and yields a multistage testing procedure for each of the statistics under consideration. See also the corresponding Section 3 of Berkes et al. (2006) for a comparison of procedures designed to discriminate between changes in the mean and long memory.

3. A SIMULATION STUDY

This section is aimed at illustrating the performance of the two-step test procedures introduced in Section 2 by means of a simulation study reporting the finite-sample properties. Note, however, that we do not intend to provide an in-depth analysis here.

SCENARIO II. To study the empirical size and power of the procedures T_n, R_n, and V_n under the change in the mean alternative $\mathcal{H}_A^{(1)}$, we have performed limited simulations for sample sizes $n = 100$, 200, 300, and 400. All results reported in this section are based on $N = 1{,}000$ repetitions. Including a variety of dependence structures, we assume that $\{Y_k\}$ is a causal AR(1) sequence defined by the difference equations

$$Y_k = \rho Y_{k-1} + \varepsilon_k, \qquad -\infty < k < \infty,$$

where $\{\varepsilon_k\}$ are independent, identically distributed standard normal random variables. The values of the autoregressive parameter ρ used in the simulations are 0.0, 0.3, 0.5, and 0.7.

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

With the exception of the case $\rho = 0.0$, the generated data are weakly correlated, which leads to the question of how to choose the bandwidths $q = q(n)$ of the Bartlett estimator. Under the null hypothesis, Andrews (1991) derived the optimal bandwidth for this estimator, which depends on the unknown spectral density of the errors. Because we are mainly concerned with data suffering from a structural break in the form of a mean change (or from random walk–type behavior), we rely instead on the recommendations given by Zhang, Gabrys, and Kokoszka (2007), who proposed a data-driven procedure that depends on the sample size, the location of the estimated change-point, and the direction of the change. They also studied other estimators for σ^2, including the vector autoregression heteroskedasticity and autocorrelation consistent (VARHAC) estimator introduced by Den Haan and Levin (1996), but their simulations showed inferior performance when compared to the Bartlett estimator. In particular, based on the extensive simulation studies in Zhang et al. (2007), we have chosen $q(100) = 4$, $q(200) = 7$, $q(300) = 9$, and $q(400) = 11$ for the first step and $q^*(100) = 3$, $q^*(200) = 4$, $q^*(300) = 5$, and $q^*(400) = 6$ for the second step of the procedures.

For the simulations, we chose the values $\Delta = 0$, 1, 1.5, and 2 as sizes of the change that occurs at time $k^* = n/2$. The results are documented in Tables 1–3

TABLE 1. Simulated rejection probabilities for T_n/s_n under $\mathcal{H}_0$ in the first step and under $\mathcal{H}_A^{(1)}$ in the second step. The change in the mean occurs at $k^* = n/2$, the size of the change is Δ, and the innovations are AR(1) processes with parameter ρ. The level is 0.1 in both steps.

T_n/s_n		$n = 100$		$n = 200$		$n = 300$		$n = 400$	
ρ	Δ	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$
0.0	0.0	0.060	0.000	0.075	0.004	0.072	0.111	0.076	0.026
0.0	1.0	0.995	0.023	1.000	0.039	1.000	0.065	1.000	0.060
0.0	1.5	1.000	0.033	1.000	0.032	1.000	0.057	1.000	0.066
0.0	2.0	1.000	0.024	1.000	0.055	1.000	0.058	1.000	0.071
0.3	0.0	0.076	0.040	0.108	0.037	0.072	0.097	0.119	0.118
0.3	1.0	0.875	0.023	0.996	0.056	1.000	0.078	1.000	0.081
0.3	1.5	0.998	0.033	1.000	0.059	1.000	0.074	1.000	0.073
0.3	2.0	1.000	0.004	1.000	0.084	1.000	0.082	1.000	0.084
0.5	0.0	0.109	0.073	0.101	0.069	0.113	0.080	0.093	0.082
0.5	1.0	0.691	0.041	0.932	0.087	0.980	0.102	0.996	0.094
0.5	1.5	0.954	0.049	1.000	0.080	1.000	0.128	1.000	0.082
0.5	2.0	0.997	0.060	1.000	0.113	1.000	0.109	1.000	0.111
0.7	0.0	0.151	0.046	0.162	0.117	0.154	0.182	0.118	0.102
0.7	1.0	0.466	0.131	0.659	0.185	0.798	0.160	0.873	0.144
0.7	1.5	0.729	0.095	0.911	0.170	0.987	0.180	0.997	0.156
0.7	2.0	0.898	0.119	0.992	0.213	1.000	0.201	1.000	0.180

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

TABLE 2. Simulated rejection probabilities for R_n/s_n under $\mathcal{H}_0$ in the first step and under $\mathcal{H}_A^{(1)}$ in the second step. The change in the mean occurs at $k^* = n/2$, the size of the change is Δ, and the innovations are AR(1) processes with parameter ρ. The level is 0.1 in both steps.

R_n/s_n		$n = 100$		$n = 200$		$n = 300$		$n = 400$	
ρ	Δ	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$
0.0	0.0	0.029	0.000	0.032	0.031	0.049	0.041	0.046	0.044
0.0	1.0	0.856	0.048	1.000	0.007	1.000	0.024	1.000	0.028
0.0	1.5	0.999	0.000	1.000	0.012	1.000	0.024	1.000	0.029
0.0	2.0	1.000	0.010	1.000	0.015	1.000	0.019	1.000	0.029
0.3	0.0	0.017	0.000	0.042	0.024	0.049	0.041	0.079	0.063
0.3	1.0	0.518	0.000	0.954	0.015	0.998	0.027	1.000	0.043
0.3	1.5	0.904	0.001	1.000	0.015	1.000	0.038	1.000	0.032
0.3	2.0	0.996	0.000	1.000	0.014	1.000	0.027	1.000	0.041
0.5	0.0	0.036	0.000	0.058	0.069	0.070	0.071	0.061	0.108
0.5	1.0	0.286	0.000	0.760	0.028	0.927	0.039	0.984	0.036
0.5	1.5	0.662	0.002	0.991	0.020	0.999	0.051	1.000	0.057
0.5	2.0	0.896	0.011	0.999	0.024	1.000	0.048	1.000	0.063
0.7	0.0	0.009	0.046	0.162	0.117	0.154	0.182	0.118	0.102
0.7	1.0	0.179	0.000	0.443	0.059	0.604	0.094	0.700	0.097
0.7	1.5	0.350	0.000	0.751	0.063	0.918	0.092	0.970	0.092
0.7	2.0	0.579	0.002	0.947	0.071	0.994	0.107	1.000	0.108

using a significance level $\alpha = 0.1$ in each of the two steps. The corresponding critical values $c(\alpha)$ have been determined using the tables in Shorack and Wellner (1986). The distributions of the limits in (2.12) and (2.13) are tabulated in Tables 1 and 2 of Shorack and Wellner (1986, pp. 143, 144), whereas the critical values for the limit distribution in (2.14) can be computed from Table 1 using the formula for the limit distribution of the Watson statistic on page 148 of Shorack and Wellner (1986). The precise numerical values are

$$c_T(0.1) = 1.225, \qquad c_R(0.1) = 1.615, \quad \text{and} \quad c_V(0.1) = 0.152$$

for the first step of the respective procedures and

$$c_T^*(0.1) = 1.350, \qquad c_R^*(0.1) = 1.740, \quad \text{and} \quad c_V^*(0.1) = 0.185$$

for the second step.

Note that in all tables the first row ($\Delta = 0$) for each of the choices of ρ gives the empirical size of the test. In the first step, we report the percentage of rejections of the null hypothesis $\mathcal{H}_0$ and, in the second step, the proportion of rejections of $\mathcal{H}_A^{(1)}$. It is apparent from Tables 1–3 that the performance of V_n is inferior to the performance of T_n and R_n in the case of a change in the mean. Although sensitive

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

TABLE 3. Simulated rejection probabilities for V_n/s_n^2 under $\mathcal{H}_0$ in the first step and under $\mathcal{H}_A^{(1)}$ in the second step. The change in the mean occurs at $k^* = n/2$, the size of the change is Δ, and the innovations are AR(1) processes with parameter ρ. The level is 0.1 in both steps.

V_n/s_n^2		$n = 100$		$n = 200$		$n = 300$		$n = 400$	
ρ	Δ	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$
0.0	0.0	0.079	0.063	0.089	0.079	0.091	0.121	0.078	0.064
0.0	1.0	0.973	0.048	1.000	0.061	1.000	0.079	1.000	0.065
0.0	1.5	1.000	0.044	1.000	0.061	1.000	0.071	1.000	0.069
0.0	2.0	1.000	0.029	1.000	0.066	1.000	0.071	1.000	0.079
0.3	0.0	0.106	0.094	0.099	0.101	0.100	0.130	0.130	0.115
0.3	1.0	0.798	0.049	0.983	0.089	0.999	0.115	1.000	0.115
0.3	1.5	0.985	0.073	1.000	0.108	1.000	0.101	1.000	0.101
0.3	2.0	1.000	0.092	1.000	0.107	1.000	0.115	1.000	0.116
0.5	0.0	0.167	0.102	0.151	0.126	0.121	0.165	0.120	0.133
0.5	1.0	0.642	0.111	0.895	0.159	0.964	0.145	0.990	0.126
0.5	1.5	0.904	0.126	0.998	0.163	0.999	0.136	1.000	0.147
0.5	2.0	0.977	0.119	0.999	0.172	1.000	0.167	1.000	0.185
0.7	0.0	0.300	0.280	0.162	0.318	0.214	0.248	0.147	0.279
0.7	1.0	0.543	0.265	0.664	0.304	0.757	0.299	0.819	0.255
0.7	1.5	0.713	0.279	0.895	0.336	0.973	0.297	0.984	0.248
0.7	2.0	0.879	0.303	0.983	0.362	0.996	0.318	1.000	0.276

to level shifts, the rescaled variance analysis therefore cannot be recommended under Scenario II for the given (moderate) sample sizes. A comparison of T_n and R_n reveals that the latter test statistic is more conservative. For the range of dependence under consideration, both the CUSUM and the R/S statistic work well. It can be seen, however, that an increase of the autoregressive parameter ρ increases the size distortion of T_n, so that larger sample sizes are required in the case of $\rho = 0.7$. The statistic R_n is much less sensitive to higher correlations. On the other hand, R_n does not detect small changes as effectively as T_n.

One source of the size distortion for large values of $|\rho|$ is the slow rate of convergence to the limit distribution in (2.2). As a consequence, the asymptotic test procedures should in this case only be applied if the sample size is reasonably large. Another reason for the distortions is the estimation of σ^2 with the Bartlett estimator. To the best of the authors' knowledge there are no definitive guidelines on how to select the bandwidth $q(n)$ in the case of structural breaks. Zhang et al. (2007) provide recommendations that have worked well for most of the simulations reported here but still leave room for improvement. Finally, the performance of the test statistics in the second step is inferior to the first step because of the

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

TABLE 4. Simulated rejection probabilities for T_n/s_n, R_n/s_n, and V_n/s_n^2 under $\mathcal{H}_0$ in the first step and under $\mathcal{H}_A^{(1)}$ in the second step when in fact the true model is given by $\mathcal{H}_A^{(2)}$. The errors are random walks based on AR(1) innovations with parameter ρ. The level is 0.1 in both steps.

		$n = 100$		$n = 200$		$n = 300$		$n = 400$	
ρ		$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$
0.0	T_n	0.965	0.871	0.878	0.930	0.918	0.929	0.942	0.952
0.0	R_n	0.813	0.172	0.881	0.772	0.940	0.918	0.948	0.956
0.0	V_n	0.938	0.828	0.962	0.959	0.982	0.988	0.979	0.991
0.3	T_n	0.846	0.618	0.893	0.888	0.909	0.926	0.940	0.943
0.3	R_n	0.806	0.185	0.913	0.802	0.936	0.924	0.942	0.963
0.3	V_n	0.929	0.817	0.959	0.963	0.909	0.984	0.940	0.990
0.5	T_n	0.798	0.650	0.902	0.898	0.908	0.923	0.938	0.958
0.5	R_n	0.706	0.191	0.894	0.830	0.926	0.923	0.963	0.957
0.5	V_n	0.901	0.872	0.949	0.976	0.962	0.981	0.980	0.989
0.7	T_n	0.819	0.690	0.892	0.906	0.929	0.937	0.961	0.966
0.7	R_n	0.707	0.198	0.916	0.836	0.942	0.941	0.978	0.959
0.7	V_n	0.914	0.890	0.961	0.975	0.975	0.990	0.990	0.990
0.9	T_n	0.848	0.763	0.909	0.926	0.938	0.963	0.968	0.979
0.9	R_n	0.819	0.303	0.940	0.915	0.954	0.970	0.974	0.985
0.9	V_n	0.949	0.944	0.975	0.987	0.985	0.994	0.990	1.000

additional estimation of the change-point k^*, which incorporates more randomness into the procedures. It turns out, however, that neither the degree of dependence nor the variance estimation significantly influences the power of the tests against random walk alternatives. This will be considered next.

SCENARIO III. In the second part of the simulation study we turn our attention to the alternative $\mathcal{H}_A^{(2)}$. We have simulated random walks $\{Y_k\}$ based on AR(1) innovations with various parameters. The results are presented in Table 4. All three test statistics have excellent power for the sample sizes $n = 300$ and $n = 400$. For the smaller sample sizes the discriminating power between $\mathcal{H}_A^{(1)}$ and $\mathcal{H}_A^{(2)}$ is still very good for V_n, but R_n performs significantly worse, in particular in the second step. The CUSUM statistic T_n performs reasonably well for the sample size $n = 100$.

There is an interesting loss in power for moderate values of ρ that disappears as the sample size increases. All test statistics work well whether the data exhibit no or little dependence or a high degree of persistence. For $\rho = 0.5$ and $n = 100$, however, T_n, R_n, and V_n all simultaneously lack power, which is most visible for the rescaled range analysis.

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

TABLE 5. Simulated rejection probabilities for T_n/s_n, R_n/s_n, and V_n/s_n^2 under $\mathcal{H}_0$ in the first step and under $\mathcal{H}_A^{(1)}$ in the second step when in fact the true model is given by $\mathcal{H}_A^{(3)}$. The errors are AR(1) processes with parameter ρ until $k^* = n/2$; then they change into random walks based on the same AR(1) process. The level is 0.1 in both steps.

		$n = 100$		$n = 200$		$n = 300$		$n = 400$	
ρ		$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$	$\mathcal{H}_0$	$\mathcal{H}_A^{(1)}$
0.0	T_n	0.847	0.422	0.905	0.741	0.913	0.824	0.933	0.880
0.0	R_n	0.678	0.028	0.882	0.472	0.926	0.719	0.948	0.829
0.0	V_n	0.901	0.572	0.948	0.845	0.956	0.917	0.964	0.934
0.3	T_n	0.840	0.416	0.894	0.735	0.938	0.850	0.939	0.878
0.3	R_n	0.675	0.025	0.869	0.480	0.934	0.730	0.939	0.807
0.3	V_n	0.885	0.624	0.938	0.859	0.969	0.903	0.967	0.939
0.5	T_n	0.835	0.435	0.916	0.783	0.933	0.833	0.941	0.894
0.5	R_n	0.682	0.021	0.886	0.488	0.935	0.737	0.944	0.834
0.5	V_n	0.901	0.630	0.956	0.849	0.960	0.909	0.962	0.944
0.7	T_n	0.843	0.508	0.911	0.787	0.927	0.855	0.931	0.881
0.7	R_n	0.716	0.039	0.909	0.540	0.936	0.770	0.945	0.849
0.7	V_n	0.931	0.728	0.953	0.893	0.964	0.932	0.958	0.951
0.9	T_n	0.889	0.597	0.920	0.865	0.942	0.913	0.960	0.937
0.9	R_n	0.791	0.051	0.916	0.663	0.952	0.850	0.959	0.912
0.9	V_n	0.961	0.809	0.967	0.952	0.973	0.968	0.984	0.981

SCENARIO IV. In the last part of the simulation study, we focus on weakly stationary innovations (modeled again by the AR(1) processes from before) that turn into a random walk at time $k^* = n/2$. The results of the Monte Carlo experiment can be found in Table 5. In principle, we are able to draw the same conclusions as in the case of the second alternative $H_A^{(2)}$. However, because of the change in persistence in the middle of the data, the discriminating power is less under $H_A^{(3)}$ in the second step of the procedures, visible in particular for the smaller sample sizes. These differences have almost disappeared for $n = 400$.

4. AN APPLICATION TO STOCK RETURN DATA

We consider daily returns from the German stock market index DAX between 1/2/92 and 12/30/99. In this observation period we have a total of $n = 2{,}009$ observations. We work with the log returns $100\log(P_k/P_{k-1})$, where P_k denotes the index value at time k. The log returns are displayed in Figure 1. Judging from the graph, there appear to be larger fluctuations toward the end of the observation period that may be caused by turning the previously stationary data into a random

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

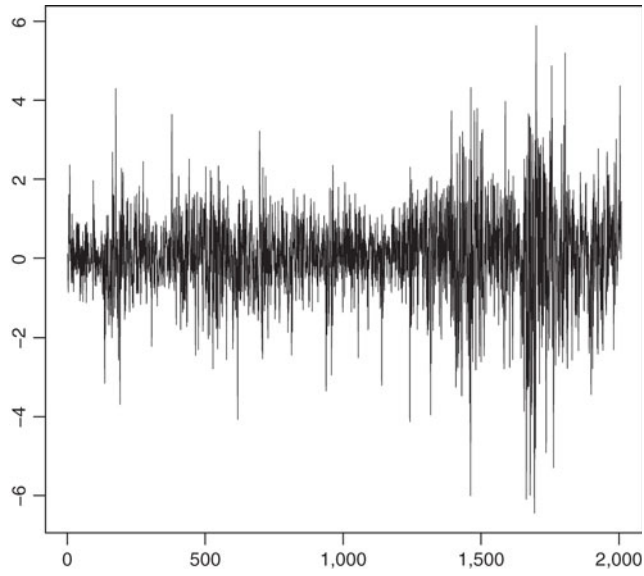


FIGURE 1. Log returns, DAX 30 (1/2/92–12/30/99).

walk. At first, we hence test the null hypothesis H_0 against the alternative of a structural break. We focus here only on the procedure based on the CUSUM statistic T_n, because applications of the rescaled range and variance procedures yield qualitatively similar results. Let

$$Z_n(k) = \frac{1}{\sqrt{ns_n^2}} \left| \sum_{i=1}^{k} (X_i - \bar{X}_n) \right|, \qquad 1 \le k \le n,$$

and note that $T_n/s_n = \max_k Z_n(k)$. The graph of $Z_n(k)$ in Figure 2 shows that the null hypothesis of a constant mean and stationary errors cannot be rejected.

In Figure 3, we display the squared log returns and, in Figure 4, the graph of $Z_n(k)$ for these squared log returns. The maximum of the test statistic is reached at $\hat{k} = 1,385$, and it is significant at the 5% level (the corresponding critical value is indicated in the graph by a horizontal line). According to our theory, we split the data into two subsets with time lags k ranging between 1 and 1,385 and 1,386 and 2,008, respectively. Figure 5 contains the sample paths of

$$Z_n^*(k) = \begin{cases} \dfrac{1}{\sqrt{\hat{k}s_{n,1}^2}} \left| \displaystyle\sum_{i=1}^{k} (X_i - \bar{X}_{\hat{k}}) \right|, & 1 \le k \le \hat{k}, \\ \dfrac{1}{\sqrt{(n-\hat{k})s_{n,2}^2}} \left| \displaystyle\sum_{i=\hat{k}+1}^{k} (X_i - \tilde{X}_{\hat{k}}) \right|, & \hat{k} < k \le n, \end{cases}$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

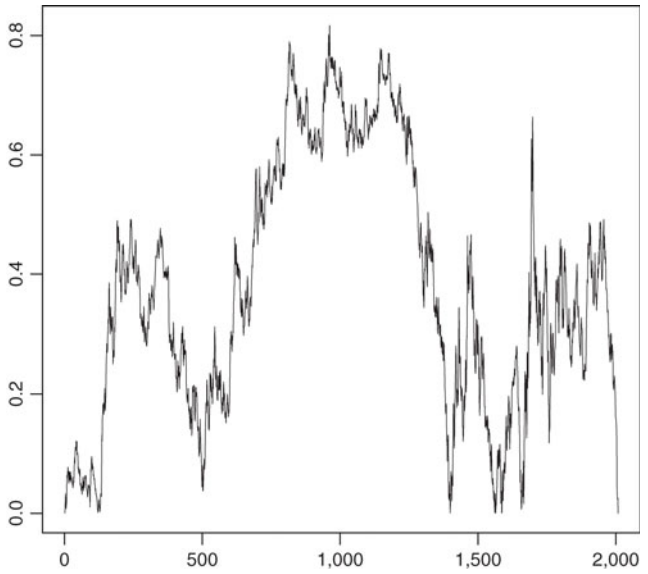


FIGURE 2. The graph of $Z_{2,008}(k)$.

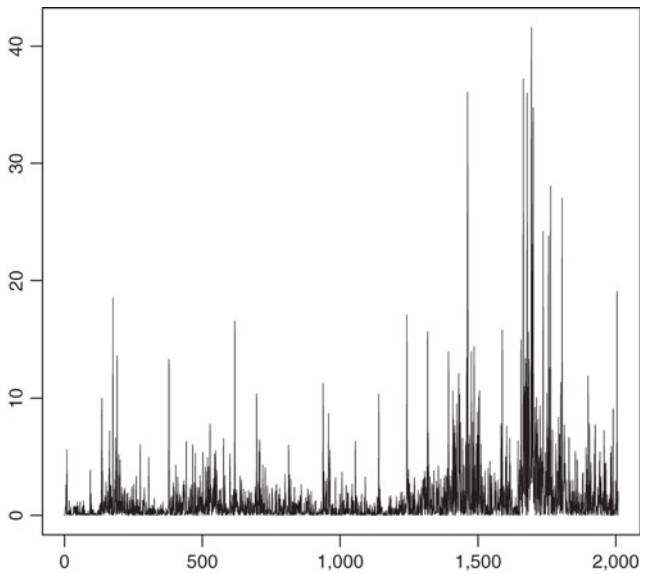


FIGURE 3. Squared log returns, DAX 30 (1/2/92–12/30/99).

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

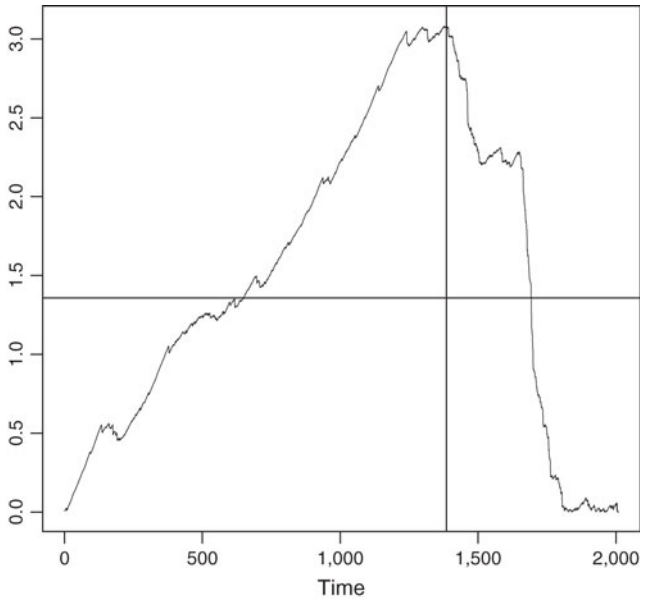


FIGURE 4. The graph of $Z_{2,008}(k)$ for the squared log returns.

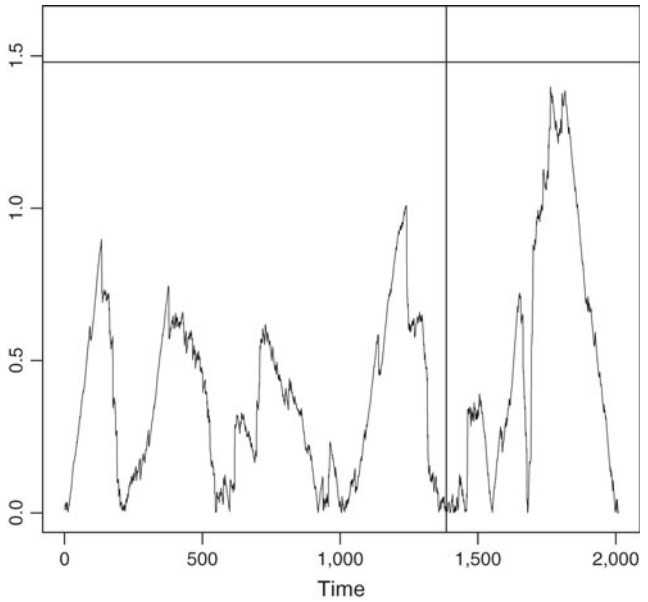


FIGURE 5. The graph of $Z^*_{2008}(k)$ for the squared log returns.

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

The plot reveals that there are no further changes in the expected squared log returns in the resulting subsets. Hence, the behavior of the log returns can be explained with weakly dependent errors having constant means but a volatility increase after observation 1,385.

5. PROOFS

Proof of Theorem 2.1. By (2.1) and (2.2) we have, with $\bar{Y}_n = 1/n \sum_{i=1}^n Y_i$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} (X_i - \bar{X}_n) = \frac{1}{\sqrt{m}} \sum_{i=1}^{\lfloor nt \rfloor} (Y_i - \bar{Y}_n) \overset{\mathcal{D}[0,1]}{\to} \sigma B(t) \qquad (n \to \infty),$$

where the process $\{B(t) : t \in [0,1]\}$, defined for $t \in [0,1]$ by $B(t) = W(t) - tW(1)$, is a standard Brownian bridge. Thus, the continuous mapping theorem and condition (2.11) imply (2.12)–(2.14). ■

Proof of Theorem 2.2. Berkes et al. (2006) proved (2.19). Following their arguments, the asymptotic consistency results of (2.20) and (2.21) are established in a similar fashion. ■

LEMMA 5.1. *If conditions (2.5) and (2.9) are satisfied, then, as* $n \to \infty$,

$$\left(\frac{1}{\sqrt{n}} Y_{\lfloor nt \rfloor}, \frac{S_n^2(t)}{n^2 q(\lfloor nt \rfloor)} \right) \overset{\mathcal{D}^2[0,1]}{\to} \left(\sigma W(t), \sigma^2 \int_0^t \left(W(x) - \frac{1}{t} \int_0^t W(u)\, du \right)^2 dx \right),$$

where $\{W(t) : t \in [0,1]\}$ *denotes a Brownian motion and, for* $t \in [0,1]$, $S_n^2(t) = n s_{\lfloor nt \rfloor}^2$.

Proof. Using $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$, for $t \in [0,1]$, rewrite

$$S_n^2(t) = \sum_{i=1}^{\lfloor nt \rfloor} (Y_i - \bar{Y}_{\lfloor nt \rfloor})^2 + 2 \sum_{j=1}^{q(\lfloor nt \rfloor)} w_j(q(\lfloor nt \rfloor)) \sum_{i=1}^{\lfloor nt \rfloor - j} (Y_i - \bar{Y}_{\lfloor nt \rfloor})(Y_{i+j} - \bar{Y}_{\lfloor nt \rfloor})$$

$$= B_{1,n}(t) + B_{2,n}(t).$$

First we establish

$$\sup_{0<t<1} \frac{|B_{1,n}(t)|}{n^2 q(\lfloor nt \rfloor)} = o_P(1) \qquad (n \to \infty). \tag{5.1}$$

To this end, observe that, for any $\varepsilon > 0$,

$$\sup_{0<t\le\varepsilon} \frac{1}{\lfloor nt \rfloor} \left(\sum_{i=1}^{\lfloor nt \rfloor} Y_i \right)^2 \le \lfloor n\varepsilon \rfloor \sup_{0\le t\le\varepsilon} Y_{\lfloor nt \rfloor}^2$$

and, by condition (2.5),

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

$$\frac{\lfloor n\varepsilon \rfloor}{n^2} \sup_{0\le t\le \varepsilon} Y^2_{\lfloor nt \rfloor} \xrightarrow{\mathcal{D}} \sigma^2 \varepsilon \sup_{0\le t\le \varepsilon} W^2(t) \qquad (n\to\infty).$$

By assumption, $\inf_{0\le t\le \varepsilon} q(\lfloor nt \rfloor) > 0$. Therefore, for any $x > 0$,

$$\lim_{\varepsilon\to 0} \limsup_{n\to\infty} P\left\{ \sup_{0\le t\le \varepsilon} \frac{1}{n^2 q(\lfloor nt \rfloor)} \frac{1}{\lfloor nt \rfloor} \left(\sum_{i=1}^{\lfloor nt \rfloor} Y_i \right)^2 \ge x \right\} = 0.$$

Similar arguments yield also that, for any $x > 0$,

$$\lim_{\varepsilon\to 0} \limsup_{n\to\infty} P\left\{ \sup_{0\le t\le \varepsilon} \frac{1}{n^2 q(\lfloor nt \rfloor)} \sum_{i=1}^{\lfloor nt \rfloor} Y_i^2 \ge x \right\} = 0.$$

On applying condition (2.5) again, it holds that, for any $\varepsilon > 0$ as $n\to\infty$,

$$\frac{B_{1,n}(t)}{n^2\sigma^2} = \frac{1}{n^2\sigma^2} \left[\sum_{i=1}^{\lfloor nt \rfloor} Y_i^2 - \frac{1}{\lfloor nt \rfloor} \left(\sum_{i=1}^{\lfloor nt \rfloor} Y_i \right)^2 \right]$$

$$\times \xrightarrow{\mathcal{D}[0,1]} \int_0^t W^2(x)\,dx - \frac{1}{t} \left(\int_0^1 W(s)\,ds \right)^2.$$

Because $\inf_{\varepsilon\le t\le 1} q(\lfloor nt \rfloor) \to \infty$ as $n\to\infty$, the proof of (5.1) is complete.

Applying the extension theorem in Berkes and Philipp (1979), we can and shall assume without loss of generality that all random processes introduced so far and in what follows are defined on the same probability space. Combining assumption (2.5) with the Dudley–Skorohod–Wichura representation theorem (see Shorack and Wellner, 1986, p. 47), there is a sequence of Brownian motions $\{W_n(t) : t \ge 0\}$ such that

$$\max_{1\le i\le n} |Y_i - \sigma W_n(i)| = o_P\left(\sqrt{n}\right) \qquad (n\to\infty). \tag{5.2}$$

On setting $\bar{W}_n(\lfloor nt \rfloor) = \frac{1}{\lfloor nt \rfloor} \sum_{i=1}^{\lfloor nt \rfloor} W_n(i)$, the latter relation (5.2) yields for the term $B_{2,n}(t)$, as $n\to\infty$,

$$\sup_{0\le t\le 1} \left| \sum_{j=1}^{q(\lfloor nt \rfloor)} w_j(q(\lfloor nt \rfloor)) \sum_{i=1}^{\lfloor nt \rfloor - j} (Y_i - \bar{Y}_{\lfloor nt \rfloor})(Y_{i+j} - \bar{Y}_{\lfloor nt \rfloor}) \right.$$

$$\left. -\sigma^2 \sum_{j=1}^{q(\lfloor nt \rfloor)} w_j(q(\lfloor nt \rfloor)) \sum_{i=1}^{\lfloor nt \rfloor - j} (W_n(i) - \bar{W}_n(\lfloor nt \rfloor))(W_n(i+j) - \bar{W}_n(\lfloor nt \rfloor)) \right|$$

$$\le C \left(\sup_{0\le t\le 1} \lfloor nt \rfloor q(\lfloor nt \rfloor) \right) \left(\max_{1\le i\le n} |\sigma W_n(i) - Y_i| \right)$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

$$\times \left(\sigma \sup_{0\le t\le n} |W_n(t)| + \max_{1\le i\le n} |\sigma W_n(i) - Y_i| \right)$$

$$= \mathcal{O}_P(n^2) \sup_{0\le t\le 1} q(\lfloor nt \rfloor),$$

because $\sup_{0\le t\le n} |W_n(t)| = \mathcal{O}_P(\sqrt{n})$. Thus we have

$$\sup_{0\le t\le 1} \left| \frac{B_{2,n}(t)}{n^2 q(\lfloor nt \rfloor)} - \frac{\sigma^2 B_{3,n}(t)}{n^2 q(\lfloor nt \rfloor)} \right| = o_P(1) \qquad (n \to \infty),$$

where

$$B_{3,n}(t) = \sum_{j=1}^{q(\lfloor nt \rfloor)} w_j(q(\lfloor nt \rfloor)) \sum_{i=1}^{\lfloor nt \rfloor - j} (W(i) - \bar{W}_n(\lfloor nt \rfloor))(W_n(i+j) - \bar{W}_n(\lfloor nt \rfloor)).$$

Note that we have suppressed the index n when dealing with the Wiener processes because the distribution of $W_n(i)$ does not depend on it. Now the modulus of continuity of Brownian motion (see Csörgő and Révész, 1981) implies, as $n \to \infty$,

$$\sup_{0\le t\le 1} \frac{1}{n^2 q(\lfloor nt \rfloor)} \left| B_{3,n}(t) - \sum_{j=1}^{q(\lfloor nt \rfloor)} w_j(q(\lfloor nt \rfloor)) \sum_{i=1}^{\lfloor nt \rfloor} (W(i) - \bar{W}_n(\lfloor nt \rfloor))^2 \right|$$

$$= \mathcal{O}_P\left(\sqrt{\frac{\log n}{n}} \sup_{1\le i\le n} q(i) \right) = o_P(1).$$

It is easy to see that, as $n \to \infty$,

$$\frac{2}{n} \sum_{j=1}^{q(\lfloor nt \rfloor)} w_j(q(\lfloor nt \rfloor)) \sum_{i=1}^{\lfloor nt \rfloor} (W(i) - \bar{W}_n(\lfloor nt \rfloor))^2$$

$$\xrightarrow{\mathcal{D}[0,1]} \int_0^t \left(W(x) - \frac{1}{t} \int_0^t W(u)\, du \right)^2 dx,$$

so that the assertion of Lemma 5.1 readily follows. ■

Proof of Theorem 2.3. The results are established after a joint application of Lemma 5.1 and the continuous mapping theorem. ■

Proof of Theorem 2.4. Because $\max_{1\le i\le k^*} |\sum_{1\le j\le i} Y_j| = \mathcal{O}_P(n^{1/2})$, the first k^* Y's play no role in the limit. Hence one can easily modify the proofs of Lemma 5.1 to establish Theorem 2.4. ■

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

LEMMA 5.2. *If the conditions of Theorem 2.5 are satisfied, then, as* $n \to \infty$,

$$\left(\frac{1}{\sqrt{\hat{k}}} \sum_{i=1}^{\lfloor nt \rfloor} (X_i - \bar{X}_{\hat{k}}) I\{\lfloor nt \rfloor \le \hat{k}\}, \frac{1}{\sqrt{n-\hat{k}}} \sum_{i=\hat{k}+1}^{\lfloor nt \rfloor} (X_i - \tilde{X}_{\hat{k}}) I\{\lfloor nt \rfloor > \hat{k}\} \right)$$

$$\stackrel{\mathcal{D}[0,1]}{\to} \left(\frac{\sigma}{\sqrt{\theta}} \left[W_{(1)}(t) - \frac{t}{\theta} W_{(1)}(\theta) \right] I\{t \le \theta\}, \right.$$

$$\left. \frac{\sigma}{\sqrt{1-\theta}} \left[W_{(2)}(t-\theta) - \frac{t-\theta}{1-\theta} W_{(2)}(1-\theta) \right] I\{t > \theta\} \right),$$

where $W_{(1)}(t)$ *and* $W_{(2)}(t)$ *are independent Brownian motions.*

Proof. On combining (2.2) with Skorohod–Dudley–Wichura representation, we can define a sequence of Brownian motions $\{W_n(t) : t \ge 0\}$ such that

$$\sup_{0 \le t \le 1} \left| \sum_{i=1}^{\lfloor nt \rfloor} Y_i - \sigma W_n(nt) \right| = o_P\left(\sqrt{n}\right) \qquad (n \to \infty). \tag{5.3}$$

Condition (2.3) yields

$$\max_{1 \le k \le \hat{k}} \left| \sum_{i=1}^{k} (X_i - \bar{X}_{\hat{k}}) - \sum_{i=1}^{k} (Y_i - \bar{Y}_{\hat{k}}) \right| \le 2|\hat{k} - k^*|\Delta = o_P\left(\sqrt{n}\right) \qquad (n \to \infty), \tag{5.4}$$

where $\bar{Y}_{\hat{k}} = \frac{1}{\hat{k}} \sum_{i=1}^{\hat{k}} Y_i$. Using (5.3), we get that, as $n \to \infty$,

$$\max_{1 \le k \le \hat{k}} \left| \sum_{i=1}^{k} (Y_i - \bar{Y}_{\hat{k}}) - \sigma \left(W_n(k) - \frac{\hat{k}}{n} W_n(\hat{k}) \right) \right| = o_P\left(\sqrt{n}\right).$$

Note that (2.16) and (2.17) imply

$$\frac{\hat{k}}{n} \stackrel{P}{\to} \theta \qquad (n \to \infty). \tag{5.5}$$

Therefore, the almost sure continuity of $\{W_n(t)\}$ also gives

$$\left| \frac{n}{\hat{k}} W_n(\hat{k}) - \frac{1}{\theta} W_n(\theta n) \right| = o_P\left(\sqrt{n}\right) \qquad (n \to \infty). \tag{5.6}$$

Next, we write

$$\max_{1 \le k \le n} \left| \left(W_n(k) - \frac{k}{\hat{k}} W_n(\hat{k}) \right) I\{k \le \hat{k}\} - \left(W_n(k) - \frac{k}{n\theta} W_n(\theta n) \right) I\{k \le n\theta\} \right|$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

$$\leq \max_{1\leq k\leq n}\left|\left[\frac{k}{\hat{k}}-\frac{k}{n\theta}\right]W_n(\hat{k})\right|+\max_{1\leq k\leq n}\left|\frac{k}{n\theta}\left(W_n(\hat{k})-W_n(n\theta)\right)\right|$$

$$+\max_{\min(\hat{k},n\theta)\leq k\leq\max(\hat{k},n\theta)}\left|W_n(k)-\frac{k}{n\theta}W_n(n\theta)\right|$$

$$=A_{n,1}+A_{n,2}+A_{n,3}.$$

Clearly, $|W_n(\hat{k})|\leq\sup_{0\leq t\leq 1}|W_n(t)|=\mathcal{O}_P(n^{1/2})$, and so by (5.5) we have

$$|A_{n,1}|=o_P(n^{1/2}).$$

Also, if W denotes a Brownian motion, then

$$n^{-1/2}\sup_{|h|\leq\epsilon}|W(n\theta)-W(n(\theta+\epsilon))|\stackrel{\mathcal{D}}{=}\sup_{|h|\leq\epsilon}|W(\theta)-W(\theta+h)|\to 0\quad\text{a.s.}$$

as $\epsilon\to 0$. Thus (5.5) yields

$$|A_{n,2}|=o_P(n^{1/2}).$$

The scale transformation of the Brownian motion W gives

$$n^{-1/2}\sup_{n(\theta-\epsilon)\leq x\leq n(\theta+\epsilon)}\left|W(x)-\frac{x}{n\theta}W(n\theta)\right|$$

$$\stackrel{\mathcal{D}}{=}\sup_{\theta-\epsilon\leq t\leq\theta+\epsilon}\left|W(t)-\frac{t}{\theta}W(\theta)\right|$$

$$\stackrel{\mathcal{D}}{=}\theta^{1/2}\sup_{1-\epsilon/\theta\leq s\leq 1+\epsilon/\theta}|W(s)-sW(1)|\to 0\quad\text{a.s.}$$

as $\epsilon\to 0$ because of the almost sure continuity of W at 1. Again using (5.5) we conclude that

$$|A_{n,3}|=o_P(n^{1/2}).$$

Hence, we get

$$\max_{1\leq k\leq n}\left|\sum_{i=1}^{\lfloor nt\rfloor}(X_i-\bar{X}_{\hat{k}})I\{\lfloor nt\rfloor\leq\hat{k}\}-\sigma\left(W_n(nt)-\frac{t}{\theta}W_n(n\theta)\right)I\{t\leq\theta\}\right|=o_P\left(\sqrt{n}\right).$$

Similar arguments also imply

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

$$\max_{1\le k\le n}\left|\sum_{i=1}^{\lfloor nt\rfloor}(X_i-\tilde{X}_{\hat{k}})I\{\lfloor nt\rfloor>\hat{k}\}\right.$$

$$\left.-\sigma\left(W_n(nt)-W_n(n\theta)-\frac{t-\theta}{1-\theta}[W_n(n)-W_n(n\theta)]\right)I\{t>\theta\}\right|=o_P\left(\sqrt{n}\right).$$

Because Brownian motions have independent increments,

$$\left\{W_n(nt)-\frac{t}{\theta}W_n(t):t\in[0,\theta]\right\},$$

$$\left\{W_n(nt)-W_n(n\theta)-\frac{t-\theta}{1-\theta}[W_n(n)-W_n(n\theta)]:t\in(\theta,1]\right\}$$

are independent. Also, the self-similarity of $W_n(t)$ gives that $\{n^{-1/2}W_n(nt):t\in[0,\theta]\}\overset{\mathcal{D}}{=}\{W(t):t\in[0,\theta]\}$ and $\{n^{-1/2}[W_n(nt)-W_n(n\theta)]:t\in(\theta,1]\}\overset{\mathcal{D}}{=}\{W(t-\theta):t\in(\theta,1]\}$, where $\{W(t):t\in[0,1]\}$ is a Brownian motion. The proof is hence complete. ■

Proof of Theorem 2.5. The result of Lemma 5.2 implies that

$$(T_{n,1},T_{n,2})\xrightarrow{\mathcal{D}}\sigma\left(\sup_{0\le t\le\theta}\frac{1}{\sqrt{\theta}}\left|W_{(1)}(t)-\frac{t}{\theta}W_{(1)}(\theta)\right|,\right.$$

$$\left.\sup_{\theta\le t\le1}\frac{1}{\sqrt{1-\theta}}\left|W_{(2)}(t-\theta)-\frac{t-\theta}{1-\theta}W_{(2)}(1-\theta)\right|\right).$$

Hence, the asymptotic independence of $T_{n,1}$ and $T_{n,2}$ is established. By the scale transformation of $W_{(1)}$ we have

$$\sup_{0\le t\le\theta}\frac{1}{\sqrt{\theta}}\left|W_{(1)}(t)-\frac{t}{\theta}W_{(1)}(\theta)\right|=\sup_{0\le s\le1}\frac{1}{\sqrt{\theta}}|W_{(1)}(\theta s)-sW_{(1)}(\theta)|$$

$$\overset{\mathcal{D}}{=}\sup_{0\le s\le1}|W_{(1)}(s)-sW_{(1)}(1)|=\sup_{0\le s\le1}|B^{(1)}(s)|,$$

where $B^{(1)}(s)=W_{(1)}(s)-sW_{(1)}(1)$ defines a Brownian bridge. Similarly,

$$\sup_{\theta\le t\le1}\frac{1}{\sqrt{1-\theta}}\left|W_{(2)}(t-\theta)-\frac{t-\theta}{1-\theta}W_{(2)}(1-\theta)\right|$$

$$=\sup_{0\le t\le1}\frac{1}{\sqrt{1-\theta}}\left|W_{(2)}(t[1-\theta])-tW_{(2)}(1-\theta)\right|$$

$$\overset{\mathcal{D}}{=}\sup_{0\le t\le1}|W_{(2)}(t)-tW_{(2)}(1)|=\sup_{0\le t\le1}|B^{(2)}(s)|,$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

where $B^{(2)}(s) = W_{(2)}(s) - s\,W_{(2)}(1)$ defines a Brownian bridge. Thus the proof of the first statement involving the CUSUM statistics T_n^* is complete.

The proof for the adjusted range statistics R_n^* follows in a similar fashion and is hence omitted here.

To prove the final statement of Theorem 2.5 note that, after an application of Lemma 5.2, we get

$$\begin{aligned}(V_{n,1}, V_{n,2}) \xrightarrow{\mathcal{D}} \sigma^2 \Bigg(& \frac{1}{\theta}\int_0^\theta \left[\frac{1}{\sqrt{\theta}}\left(W_1(t) - \frac{t}{\theta}W_1(\theta)\right)\right]^2 dt \\ & - \frac{1}{\theta}\left[\int_0^\theta \frac{1}{\sqrt{\theta}}\left(W_1(t) - \frac{t}{\theta}W_1(\theta)\right) dt\right]^2, \\ & \frac{1}{1-\theta}\int_\theta^1 \left[\frac{1}{\sqrt{1-\theta}}\left(W_2(t-\theta) - \frac{t-\theta}{1-\theta}W_2(1-\theta)\right)\right]^2 dt \\ & - \frac{1}{1-\theta}\left[\int_\theta^1 \frac{1}{\sqrt{1-\theta}}\left(W_2(t-\theta) - \frac{t-\theta}{1-\theta}W_2(1-\theta)\right) dt\right]^2\Bigg),\end{aligned}$$

proving the asymptotic independence. Applying the change of variable formula and the scale transformation for Brownian motions yields

$$\begin{aligned}& \frac{1}{\theta}\int_0^\theta \left[\frac{1}{\sqrt{\theta}}\left(W_{(1)}(t) - \frac{t}{\theta}W_{(1)}(\theta)\right)\right]^2 dt - \frac{1}{\theta}\left[\int_0^\theta \frac{1}{\sqrt{\theta}}\left(W_{(1)}(t) - \frac{t}{\theta}W_{(1)}(\theta)\right) dt\right]^2 \\ & \quad \stackrel{\mathcal{D}}{=} \frac{1}{\theta}\int_0^\theta \left(W_{(1)}\left(\frac{t}{\theta}\right) - \frac{t}{\theta}W_{(1)}(1)\right)^2 dt - \frac{1}{\theta}\left(\int_0^\theta \left[W_{(1)}\left(\frac{t}{\theta}\right) - \frac{t}{\theta}W_{(1)}(1)\right] dt\right)^2 \\ & \quad = \int_0^1 (B^{(1)}(t))^2\,dt - \left(\int_0^1 B^{(1)}(t)\,dt\right).\end{aligned}$$

Analogously,

$$\begin{aligned}& \frac{1}{(1-\theta)^2}\left(\int_0^1 \left[W_{(2)}(t-\theta) - \frac{t-\theta}{1-\theta}W_{(2)}(1-\theta)\right]^2 dt \right. \\ & \qquad \left. - \left[\int_\theta^1 \left(W_{(2)}(t-\theta) - \frac{t-\theta}{1-\theta}W_{(2)}(1-\theta)\right) dt\right]\right) \\ & \quad \stackrel{\mathcal{D}}{=} \int_0^1 (B^{(2)}(t))^2\,dt - \left(\int_0^1 B^{(2)}(t)\,dt\right),\end{aligned}$$

completing the proof. ■

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

Proof of Theorem 2.6. It follows from Lemma 5.1 that, as $n \to \infty$,

$$\frac{1}{n^{3/2}} \sum_{i=1}^{\lfloor nt \rfloor} (X_i - \bar{X}_n) \overset{\mathcal{D}[0,1]}{\to} \sigma \left(\int_0^t W(x)\,dx - t \int_0^1 W(x)\,dx \right).$$

Thus, by the continuous mapping theorem,

$$\frac{\hat{k}}{n} \overset{\mathcal{D}}{\to} \xi \qquad (n \to \infty), \tag{5.7}$$

where

$$\xi = \inf \left\{ t \in [0,1] : \left| \int_0^t W(x)\,dx - t \int_0^1 W(x)\,dx \right| \right.$$

$$\left. = \sup_{0 \le s \le 1} \left| \int_0^s W(x)\,dx - s \int_0^1 W(x)\,dx \right| \right\}.$$

Even more follows from the preceding arguments, namely the joint convergence of the random vector

$$\left(\frac{1}{\sqrt{n}} Y_{\lfloor nt \rfloor}, \frac{S_n(t)}{n^2 q(\lfloor nt \rfloor)}, \frac{\hat{k}}{n} \right)$$

in terms of the same Wiener process $\{W(t)\}$. Next, observe that

$$\frac{1}{\sqrt{\hat{k}}} \max_{1 \le k \le \hat{k}} \left| \sum_{i=1}^{k} (Y_i - \bar{Y}_{\hat{k}}) \right| = \frac{1}{\sqrt{\hat{k}}} \max_{1 \le k \le n\hat{k}/n} \left| \sum_{i=1}^{k} Y_i - \frac{k}{\hat{k}} \sum_{i=1}^{\hat{k}} Y_i \right|$$

$$= \sqrt{\frac{n^3}{\hat{k}}} \max_{1 \le k \le n\hat{k}/n} \frac{1}{n^{3/2}} \left| \sum_{i=1}^{k} Y_i - \frac{k}{n} \frac{n}{\hat{k}} \sum_{i=1}^{\hat{k}} Y_i \right|$$

and

$$s_{n,1}^2 = \frac{S_n^2(\hat{k}/n)}{\hat{k}} = \frac{n^2 q(\hat{k})}{\hat{k}} \frac{S_n^2(\hat{k}/n)}{n^2 q(\hat{k})}.$$

So, by the continuous mapping theorem, we obtain, as $n \to \infty$,

$$\left(\frac{1}{n^{3/2}} \max_{1 \le k \le \hat{k}} \left| \sum_{i=1}^{k} Y_i - \frac{k}{\hat{k}} \sum_{i=1}^{\hat{k}} Y_i \right|, \frac{S_n^2(\hat{k}/n)}{n^2 q(\hat{k})} \right)$$

$$\overset{\mathcal{D}}{\to} \left(\sigma \sup_{0 \le t \le \xi} \left| \int_0^t W(u)\,du - \frac{t}{\xi} \int_0^{\xi} W(u)\,du \right|, \right.$$

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

$$\sigma^2 \int_0^{\xi} \left[W(x) - \frac{1}{\xi} \int_0^{\xi} W(u)\,du \right]^2 dx \Bigg).$$

Hence, as $n \to \infty$,

$$\sqrt{\frac{q(\hat{k})}{n}} \frac{T_{n,1}}{s_{n,1}} \xrightarrow{\mathcal{D}} \left(\int_0^{\xi} \left[W(x) - \frac{1}{\xi} \int_0^{\xi} W(u)\,du \right]^2 dx \right)^{-1/2}$$
$$\times \sup_{0 \le t \le \xi} \left| \int_0^t W(u)\,du - \frac{t}{\xi} \int_0^{\xi} W(u)\,du \right|.$$

Condition (2.9) and (5.7) imply $n/q(\hat{k}) \xrightarrow{P} \infty$ and therefore

$$\frac{T_{n,1}}{s_{n,1}} \xrightarrow{P} \infty \qquad (n \to \infty),$$

resulting in $T_n^* \xrightarrow{P} \infty$. Similar arguments work for the other two statements given in Theorem 2.6, and the proof is complete. ■

Proof of Theorem 2.7. Because $\max_{1 \le i \le k^*} |\max_{1 \le j \le i} Y_j| = \mathcal{O}_P(n^{1/2})$, it can be proved along the lines of the proof of Theorem 2.6. ■

REFERENCES

Andreau, E. & A. Spanos (2003) Statistical adequacy and the testing of trend versus difference stationarity. *Econometric Reviews* 22, 217–237.

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance estimation. *Econometrica* 59, 817–858.

Aue, A., I. Berkes, & L. Horváth (2006) Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli* 12, 583–608.

Bai, J. (1994) Least squares estimation of a shift in linear processes. *Journal of Time Series Analysis* 15, 453–470.

Banerjee, A., R. Lumsdaine, & Stock, J. (1992) Recursive and sequential tests of the unit root and trend break hypothesis: Theory and international evidence. *Journal of Business & Economic Statistics* 10, 271–288.

Bartlett, M.S. (1950) Periodogram analysis of continuous spectra. *Biometrika* 37, 1–16.

Belaire-Franch, J. (2005) A proof of the power of Kim's test against stationary processes with structural breaks. *Econometric Theory* 21, 1172–1176.

Berkes, I., L. Horváth, P. Kokoszka, & Q.-M. Shao (2005) Almost sure convergence of the Bartlett estimator. *Periodica Mathematica Hungarica* 51, 11–25.

Berkes, I., L. Horváth, P. Kokoszka, & Q.-M. Shao (2006) On discriminating between long-range dependence and changes in the mean. *Annals of Statistics* 34, 1140–1165.

Berkes, I. & W. Philipp (1979) Approximation theorem for independent and weakly dependent random vectors. *Annals of Probability* 7, 29–54.

Busetti, F. & A.M.R. Taylor (2004) Tests of stationarity against a change in persistence. *Journal of Econometrics* 123, 33–66.

Breidt, F.J., N. Crato, & P. de Lima (1998) On the detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* 83, 325–348.

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

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

Csörgő, M. & L. Horváth (1997) *Limit Theorems in Change-Point Analysis.* Wiley.

Csörgő, M. & P. Révész (1981) *Strong Approximations in Probability and Statistics.* Academic Press.

Den Haan, W.J. & A. Levin (1996) Inference from Parametric and Non-parametric co-Variance Matrix Estimation Procedures. National Bureau of Economic Research Technical Working Paper 2000-11.

Feller, W. (1951) The asymptotic distribution of the range of sums of independent random variables. *Annals of Mathematical Statistics* 22, 427–432.

Giraitis, L., P. Kokoszka, & R. Leipus (2001) Testing for long memory in the presence of a general trend. *Journal of Applied Probability* 10, 1002–1024.

Giraitis, L., P. Kokoszka, R. Leipus, & G. Teyssiére (2003) Rescaled variance and related tests for long memory in volatility and levels. *Journal of Econometrics* 112, 265–294.

Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and Its Application.* Academic Press.

Harvey, D.I. & T.C. Mills (2004) Tests for stationarity in a series with endogenously determined structural change. *Oxford Bulletin of Economics and Statistics* 66, 863–894.

Hurst, H. (1951) Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* 116, 770–799.

Kim, J.-Y. (2000) Detection of change in persistence of a linear time series. *Journal of Economics* 95, 97–116.

Kim, J.-Y, J. Belaire-Franch, & R. Badillo-Amador (2002) Corringendum to "Detection of change in persistence of a linear time series." *Journal of Economics* 109, 389–392.

Kuiper, N.H. (1960) Tests concerning random points on the circle. *Proceedings of the Koninklijke Nederlandse Akademie Van Wettenschappen, Series A* 63, 38–47.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, & Y. Shin (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* **54**, 159–178.

Leybourne, S., T.-H. Kim, V. Smith, & P. Newbold (2003) Tests for a change in persistence against the null of difference-stationarity. *Econometrics Journal* 6, 291–311.

Lo, A. (1991) Long-term memory in stock market prices. *Econometrica* 59, 1279–1313.

Mandelbrot, B.B. & M.S. Taqqu (1979) Robust R/S analysis of long run serial correlation. In *42nd Session of the International Statistical Institute, Manila*, book 2, pp. 69–99.

Nelson, C.R. & C.I. Plosser (1982) Trends and random walks in macroeconomics time series: Some evidence and implications. *Journal of Monetary Economics* 10, 139–162.

Nyblom, J. & T. Mäkeläinen (1983) Comparisons of tests for the presence of random walk coefficients in a simple linear model. *Journal of the American Statistical Association* 78, 856–864.

Page, E.S. (1954) Continuous inspection schemes. *Biometrika* 41, 100–105.

Perron, P. (1989) The great crash, the oil price shock, and the unit root hypothesis. *Econometrica* 57, 1361–1401.

Perron, P. (1990) Testing for a unit root in a time series with a changing mean. *Journal of Business & Economic Statistics* 8, 153–162.

Perron, P. & T.J. Vogelsang (1992) Testing for a unit root in a time series with a changing mean: Corrections and extensions. *Journal of Business & Economic Statistics* 10, 467–470.

Philipp, W. (1986) Invariance principles for independent and weakly dependent random variables. In *Dependence in Probability and Statistics*, Progress in Probability and Statistics, 11, pp. 225–268 Birkhäuser.

Shao, Q.-M. (1993) Almost sure invariance principles for mixing sequences of random variables. *Stochastic Processes and Their Applications* 48, 319–334.

Shorack, G.R. & J.A. Wellner (1986) *Empirical Processes with Applications to Statistics.* Wiley.

Taylor, A.M.R. (2005) Fluctuation tests for change in persistance. *Oxford Bulletin of Economics and Statistics* 67, 207–229.

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press

Teverovsky, V., M.S. Taqqu, & W. Willinger (1999). Stock market prices and long-range dependence. *Finance and Stochastics* 3, 1–13.

Vostrikova, L.J. (1981) Discovery of "discord" in multidimensional random processes. *Doklady Akademii Nauk SSSR* 259, 270–274.

Watson, G.S. (1961) Goodness-of-fit tests on a circle. *Biometrika* 48, 109–114.

Xiao, Z. (2001) Testing the null hypothesis of stationarity against an autoregressive unit root alternative. *Journal of Time Series Analysis* 22, 87–105.

Yao, Y.-C. (1988). Estimating the number of change-points via Schwartz's criterion. *Statistics and Probability Letters* 6, 173–177.

Zhang, A., R. Gabrys, & P. Kokoszka (2007) Discriminating between long memory and volatility shifts. *Austrian Journal of Statistics* 36, 253–275.

https://doi.org/10.1017/S0266466608090130 Published online by Cambridge University Press