ELSEVIER

Journal of Econometrics 112 (2003) 265–294

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Rescaled variance and related tests for long memory in volatility and levels

Liudas Giraitis[a,1], Piotr Kokoszka[b,2], Remigijus Leipus[c,d,*,3], Gilles Teyssière[e]

[a] *Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, UK*
[b] *Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, UT 84322-3900, USA*
[c] *Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius 2600, Lithuania*
[d] *Institute of Mathematics and Informatics, Akademijos 4, Vilnius 2600, Lithuania*
[e] *GREQAM & CORE, Centre de la Vieille Charité, 2 rue de la Charité, F-13002 Marseille, France*

Abstract

This paper studies properties of tests for long memory for general fourth order stationary sequences. We propose a rescaled variance test based on V/S statistic which is shown to have a simpler asymptotic distribution and to achieve a somewhat better balance of size and power than Lo's (Econometrica 59 (1991) 1279) modified R/S test and the KPSS test of Kwiatkowski et al. (J. Econometrics 54 (1992) 159). We investigate theoretical performance of R/S, KPSS and V/S tests under short memory hypotheses and long memory alternatives, providing a Monte Carlo study and a brief empirical example. Assumptions of the same type are used in both short and long memory cases, covering all persistent dependence scenarios. We show that the results naturally apply and the assumptions are well adjusted to linear sequences (levels) and to squares of linear ARCH sequences (volatility).

JEL classification: C12; C15; C22

Keywords: Long memory; Modified R/S statistic; KPSS statistic; V/S statistic; Linear process; LARCH model

* Corresponding author. Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius 2600, Lithuania.

E-mail addresses: l.giraitis@lse.ac.uk (L. Giraitis), piotr@math.usu.edu (P. Kokoszka), remigijus.leipus@maf.vu.lt (R. Leipus), gilles@ehess.cnrs-mrs.fr (G. Teyssière).

[1] Supported by ESRC grant R000235892 and by ESRC grant R000238212.

[2] Partially supported by EPSRC grant GR/L/78222 at the University of Liverpool.

[3] The first draft of this paper was written while this author was a Research Associate at the University of Liverpool. Partially supported by Lithuanian Science and Studies Foundation Grant K–014.

PII: S0304-4076(02)00197-5

1. Introduction

Long memory is commonly used to describe persistent dependence between the time series observations X_k as the lag increases and, in the context of covariance stationary sequences, is typically characterized by hyperbolic decay, $k^{2d-1}(0 < d < 1/2)$, of the autocovariance function $\mathrm{Cov}(X_k, X_0)$, so that it is not absolutely summable. A series is said to have short memory if the auto-covariance function is absolutely summable. These definitions are applicable to *any* stationary sequences and we adopt them in this paper.

Various tools for detecting possible long memory in time series were proposed in the econometric and statistical literature. Geweke and Porter-Hudak (1983) introduced a semiparametric procedure, whereas, following the work of Hurst (1951), Mandelbrot and Taqqu (1979) and others, Lo (1991) developed a nonparametric R/S-type test, which has become widely applicable in empirical literature. Lobato and Robinson (1998) considered a nonparametric test for $I(0)$ against long memory alternatives.

This paper deals with particular nonparametric R/S-type tests for long memory in covariance stationary sequences. The asymptotics of the tests are established under null and alternative hypotheses. In addition to Lo's test and KPSS test of Kwiatkowski et al. (1992), we consider a new test, called *rescaled variance test*, based on a V/S statistic, which differs from KPSS statistic by correction for a mean and is shown to have a simpler asymptotic distribution and to achieve a better balance of size and power than the other two tests. To obtain the corresponding asymptotics we suggest simple assumptions nesting both short and long memory cases.

Although we deal with general stationary series, particular attention is turned to linear models and nonlinear ARCH-type models, which are popular in applications. The linear models (e.g. FARIMA-type moving averages), which are well known and have been extensively investigated, are used to model a short and long memory in levels.

Nonlinear ARCH-type models, having different dependence structure from linear models, are well-suited to model some of the regularities observed in financial data. While long memory has been shown to be present in geophysical and, more recently, in network traffic data, its presence in market data is a matter of debate. Empirical studies suggest that the returns r_k on asset prices are essentially uncorrelated, but the transforms such as the squares r_k^2 or absolute values $|r_k|$, may exhibit some form of persistent (strong) dependence. While the presence of a very small correlation in returns can be to a large extent explained by factors like bid–ask spread and non-synchronous trading, see Campbell et al. (1997), long memory in returns would be a radical departure from the random walk hypothesis and the assumption of the unpredictability of asset returns which underlines the classical asset pricing theory. The presence of strong dependence in the series of squares r_k^2 and absolute values $|r_k|$ of returns does not contradict the efficient market hypothesis, but affects the volatility estimators and thus the derivative pricing formulas relying on these estimators.

Conclusions about strong correlation in squared returns are typically reached by examining autocorrelation plots and using graphical methods without applying rigorous statistical testing procedures. Even though several attempts have been made to construct long memory ARCH-type models, the theory supporting statistical conclusions is still being developed, some earlier claims, e.g. on the existence of stationary long memory ARCH processes, eventually appear to be incorrect. Interested readers are referred to Ding and Granger (1996), Baillie et al. (1996), Robinson and Zaffaroni (1998), Giraitis et al. (2000a) (further GKL (2000)), Giraitis et al. (2000c) (further GRS (2000)) and references therein. We also note that the observed characteristics, like autocorrelations and spectral estimates, which indicate strong dependence, may in fact, be due to some forms of non-stationarity like trends or changing parameters, see e.g. Mikosch and Stărică (1999), Giraitis et al. (2001), Kirman and Teyssière (2002), Teyssière (2002).

The empirical study in the present paper shows that in both linear and ARCH cases the test based on the V/S statistic is less sensitive to the choice of the bandwidth parameter q than the test based on Lo's statistic. The search of optimal choice of q remains however a complex problem. The V/S test has practically uniformly higher power than the KPSS test and, unlike for the latter, the asymptotic distribution of the V/S statistic has a simple form so that asymptotic critical values can be found using a table of the standard Kolmogorov statistic.

In order to derive the asymptotic distributions of the test statistics, a functional limit theorems are needed. The verification of the convergence of the denominator of the three statistics is also not trivial: typically a certain condition on the fourth order cumulants (see Theorem 3.1) must be assumed. The condition we propose is naturally satisfied by linear and ARCH sequences. Alternatively, appropriate mixing assumptions could be imposed (see Davidson and de Jong, 2000; Davidson, 2002).

The main objective of the present paper is to advance the relevant theory. The related empirical studies by Kirman and Teyssière (2002), Teverovsky et al. (1999a) on R/S-type and Lobato–Robinson tests for long memory provide an additional useful guidance for practical use of these procedures. They show that the long memory test of Lobato and Robinson (1998) based on the semiparametric estimation of the long memory parameter might compete with R/S-type tests in the class of linear models. A theoretical extension of the Lagrange Multiplier type test of Lobato and Robinson (1998) to the case of nonlinear models is a difficult open problem. For a recent study in this direction see Robinson and Henry (1999).

The paper is organized as follows. The test statistics are introduced in Section 2. In Section 3 we describe the null and alternative hypotheses and establish the asymptotics of the considered statistics. As particular examples of processes satisfying these hypothesis we consider linear sequences and squares of linear ARCH observations. In Section 4, we present the results of a small simulation study which offers some insight into the finite sample performance of the tests. Final comments are given in Section 5.

2. Test statistics

This section presents three R/S-type tests for long memory. We show in the sequel by means of a simulation study that the new V/S statistic, introduced in Section 2.3, outperforms the R/S and KPSS statistics. In this section, $X_1, \ldots, X_N$ is the observed sample.

2.1. Modified R/S statistic

The rescaled range, or R/S analysis was introduced by Hurst (1951) and subsequently refined by Mandelbrot and his collaborators (see Mandelbrot and Wallis, 1969; Mandelbrot, 1972, 1975; Mandelbrot and Taqqu, 1979).

Lo (1991) introduced the *modified* R/S statistic

$$Q_N(q) = \frac{1}{\hat{s}_{N,q}} \left[\max_{1 \leqslant k \leqslant N} \sum_{j=1}^{k} (X_j - \bar{X}_N) - \min_{1 \leqslant k \leqslant N} \sum_{j=1}^{k} (X_j - \bar{X}_N) \right], \tag{2.1}$$

where $\bar{X}_N$ is the sample mean $N^{-1}\sum_{j=1}^{N} X_j$ and $\hat{s}^2_{N,q}$ is an estimator of $\sigma^2 = \sum_j \mathrm{Cov}(X_j, X_0)$ defined by

$$\hat{s}^2_{N,q} = \frac{1}{N} \sum_{j=1}^{N} (X_j - \bar{X}_N)^2 + 2 \sum_{j=1}^{q} \omega_j(q) \hat{\gamma}_j. \tag{2.2}$$

In (2.2),

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

are the Bartlett weights and the $\hat{\gamma}_j$ are the sample covariances:

$$\hat{\gamma}_j = \frac{1}{N} \sum_{i=1}^{N-j} (X_i - \bar{X}_N)(X_{i+j} - \bar{X}_N), \quad 0 \leqslant j < N. \tag{2.3}$$

The classical R/S statistic corresponds to $q = 0$, so that (2.2) does not contain the second term. Whereas the classical R/S analysis focused on estimating the limit of the ratio $\log Q_N(0)/\log N$, called Hurst coefficient, Lo (1991) proposed a statistical hypothesis testing procedure to detect long memory. It should be noted that the asymptotic distribution of the statistics $Q_N(0)$ depends strongly on the correlation structure of the data and is not asymptotically parameter free, so it cannot be used to construct a test. The second term in (2.2) was suggested by Lo (1991) in order to take into account short range dependence. He has shown that by allowing q to increase slowly with the sample size, asymptotic distribution of $Q_N \equiv Q_N(q)$ is parameter free and robust to many forms of weak dependence in the data. As pointed out by Lo (1991) other windows yielding non-negative estimators of spectral density, such as Parzen window, can be used as well. Andrews (1991) provides a data-depend rule for choosing q, which, however, has only limited applicability.

The modified R/S statistic has been extensively used to detect long memory in speculative assets. To name just a few contributions, Goetzmann (1993) and Teverovsky et al. (1999b) investigated long memory in stock returns, Crato and deLima (1994), Breidt et al. (1998)—in conditional variance of stock returns. Cheung (1993a)—in foreign-exchange rates, Cheung and Lai (1993)—in gold prices, Liu et al. (1993)—in currency futures prices. Finite sample performance of the modified R/S and other statistics used to detect long memory was examined by Monte Carlo methods by Cheung (1993b) and Hauser (1997).

2.2. KPSS statistic

Another important R/S-type test is the KPSS test introduced by Kwiatkowski et al. (1992). In the context of testing for long memory in a stationary sequence the KPSS statistic takes the form:

$$T_N = \frac{1}{\hat{s}_{N,q}^2 N^2} \sum_{k=1}^{N} \left(\sum_{j=1}^{k} (X_j - \bar{X}_N) \right)^2 \tag{2.4}$$

with $\hat{s}_{N,q}^2$ given by (2.2).

Kwiatkowski et al. (1992) introduced a KPSS test as a test of trend stationarity against a unit root alternative. Lee and Schmidt (1996) used the KPSS statistic to test for the presence of long memory in a stationary time series, whereas Lee and Amsler (1997) considered non-stationary long memory alternatives. Kwiatkowski et al. (1992) and Shin and Schmidt (1992) obtained the asymptotic distribution of the KPSS statistic under the null hypothesis of trend stationarity and weakly dependent error process. Lee and Schmidt (1996) and Lee and Amsler (1997) gave its asymptotic distribution under stationary long memory and non-stationary fractionally integrated alternatives, respectively.

2.3. V/S statistic

We propose to introduce a centering in the statistic T_N and consider the following statistic, which we call V/S or rescaled variance statistic:

$$M_N = \frac{1}{\hat{s}_{N,q}^2 N^2} \left[\sum_{k=1}^{N} \left(\sum_{j=1}^{k} (X_j - \bar{X}_N) \right)^2 - \frac{1}{N} \left(\sum_{k=1}^{N} \sum_{j=1}^{k} (X_j - \bar{X}_N) \right)^2 \right]$$

with $\hat{s}_{N,q}^2$ given by (2.2).

The V/S in the title comes from *variance/S*, as the statistic M_N has the form

$$M_N = N^{-1} \frac{\widehat{\mathrm{Var}}(S_1^*, \ldots, S_N^*)}{\hat{s}_{N,q}^2}, \tag{2.5}$$

where $S_k^* = \sum_1^k (X_j - \bar{X}_N)$ are the partial sums of the observations and $\widehat{\mathrm{Var}}(S_1^*, \ldots, S_N^*) = N^{-1} \sum_1^N (S_j^* - \bar{S}_N^*)^2$ is their sample variance. In this notation, the modified R/S statistic can be written as

$$Q_N(q) = \frac{\max_{1 \leqslant k \leqslant N} S_k^* - \min_{1 \leqslant k \leqslant N} S_k^*}{\hat{s}_{N,q}}. \tag{2.6}$$

Thus, the range of the partial sums S_k^* in (2.6) has been replaced in (2.5) by their variance, and the scaling has been accordingly readjusted. Similarly, the KPSS statistic (2.4) may now be seen as a rescaled sample second moment of the partial sums. It can be hoped that "corrected for a mean" statistic M_N will be more sensitive to "shifts in variance" (see Durbin, 1973, p. 36) than T_N and will have higher power than both the R/S and KPSS statistics against long memory in squares. This is confirmed by simulation experiments presented in Section 4. We will see in Section 3 that after a simple transformation the asymptotic distribution of M_N coincides with the limiting distribution of the standard Kolmogorov statistic, see (3.18).

It should be noted that the asymptotic distributions and mutual relationships between the test statistics considered in this paper are basically the same as for the corresponding statistics based on the empirical distribution function. Namely, the relationship between the R/S and the V/S statistic M_N is analogous to that between Kuiper's (1960) and Watson's (1961) statistic in the context of goodness-of-fit testing on the circle. While the classical Kolmogorov and von Mises statistics are well suited to real-valued observations, they are not suitable for testing the goodness-of-fit of observations on a circle. To rectify this shortcoming, Kuiper (1960) introduced the following modification of the classical Kolmogorov statistic:

$$K_N = \sup_{0 \leqslant t \leqslant 1} (\hat{F}_N(t) - F(t)) - \inf_{0 \leqslant t \leqslant 1} (\hat{F}_N(t) - F(t)),$$

where $\hat{F}_N(t)$ is the empirical distribution function of the variables $X_1, \ldots, X_N$ taking values on a circle of unit length, and $F(t)$, $0 \leqslant t \leqslant 1$ is their distribution function. Watson's statistic W_N is analogous to von Mises statistic and has the following form

$$W_N = N \left[\int_0^1 (\hat{F}_N(t) - F(t))^2 \, \mathrm{d}F(t) - \left(\int_0^1 (\hat{F}_N(t) - F(t)) \, \mathrm{d}F(t) \right)^2 \right].$$

The applicability of statistics K_N and W_N is, however, not restricted to tests on the unit circle; they have been shown to have high power when one is more interested in the discrimination against shifts in variance than against shifts in mean.

In a slightly different context, Tsay (1998) considered a test of the null hypothesis of difference stationarity against stationary and non-stationary fractionally integrated alternatives. He introduced the *modified Durbin–Watson statistic* whose asymptotic distribution under appropriate assumptions differs from that of the V/S statistic in that the Brownian bridge is replaced by the Brownian motion.

3. Asymptotic theory

We derive in this section the asymptotic distribution of the R/S-type statistics under the short memory null hypothesis (Assumption S) and the long memory alternative (Assumption L). We also consider the linear and ARCH-type models as specific examples and show that these conditions are well adjusted to the specific structure of these models.

To cover basic possible dependence structures and to investigate in full the behavior of the tests for general stationary sequences, we formulate our assumptions in the same terms both for the short and long memory. These assumptions are rather non-restrictive and include conditions on the 4th order cumulants. It should be noted that although we restrict ourselves to stationary sequences, the null hypothesis can, however, be extended to include certain non-stationary processes as in Lo (1991). Since this would lead to further technical complications in the proofs, we do not pursue it here.

Throughout the paper, for stationary sequences $\{X_k\}$, we denote

$$\mu = \mathrm{E}X_k \quad \text{and} \quad \gamma_k = \mathrm{Cov}(X_k, X_0).$$

In the following, $\{W(t),\ 0 \leqslant t \leqslant 1\}$ is the standard Wiener process (Brownian motion) and $\{W^0(t),\ 0 \leqslant t \leqslant 1\}$ is the Brownian bridge $W^0(t) = W(t) - tW(1)$. By $\Rightarrow$ we denote the weak convergence of random variables, and by $\stackrel{D[0,1]}{\longrightarrow}$ the weak convergence in the space $D[0,1]$ endowed with the Skorokhod topology (see Billingsley, 1968).

Short memory null hypothesis. Typically a stationary sequence $\{X_k\}$ is said to have short memory if

$$\sum_{j=-\infty}^{\infty} |\gamma_j| < \infty. \tag{3.1}$$

In order to investigate rigorously the asymptotic behavior of the tests we need some additional assumptions. We formulate them as

Assumption S. The fourth order stationary sequence $\{X_k\}$ satisfies (3.1); convergence

$$N^{-1/2} \sum_{j=1}^{[Nt]} (X_j - \mathrm{E}X_j) \stackrel{D[0,1]}{\longrightarrow} \sigma W(t), \quad \text{as } N \to \infty \tag{3.2}$$

holds, where $\sigma^2 = \sum_{j=-\infty}^{\infty} \gamma_j \neq 0$ and the fourth order cumulants $\kappa(h,r,s)$ satisfy

$$\sup_h \sum_{r,s=-\infty}^{\infty} |\kappa(h,r,s)| < \infty. \tag{3.3}$$

Recall that

$$\kappa(h,r,s) = \mathrm{E}[(X_k - \mu)(X_{k+h} - \mu)(X_{k+r} - \mu)(X_{k+s} - \mu)] \\ -(\gamma_h\gamma_{r-s} + \gamma_r\gamma_{h-s} + \gamma_s\gamma_{h-r}). \tag{3.4}$$

Long memory alternative. We say that a stationary sequence $\{X_k\}$ has long memory (LM) if $\sum_{j=-\infty}^{\infty} |\gamma_k| = \infty$. To characterize a class of long memory sequences for which consistency of the tests can be showed and asymptotic distribution derived, we introduce a set of long memory conditions (Assumption L) using the same concepts as in the short memory case. In the following, $W_H(t)$ stands for the fractional Brownian motion with parameter H, i.e. a Gaussian process with mean zero and covariances $\mathrm{E}W_H(t_1)W_H(t_2) = 1/2\ (t_1^{2H} + t_2^{2H} - |t_1 - t_2|^{2H})$.

Assumption L. The fourth order stationary sequence $\{X_k\}$ satisfies the following conditions:

$$\gamma_k \sim ck^{2d-1}, \tag{3.5}$$

where $c > 0$ and $0 < d < 1/2$; convergence

$$N^{-1/2-d} \sum_{j=1}^{[Nt]} (X_j - \mathrm{E}X_j) \stackrel{D[0,1]}{\longrightarrow} c_d W_{1/2+d}(t), \quad \text{as } N \to \infty \tag{3.6}$$

holds, where c_d is a positive number; the cumulants $\kappa(h,r,s)$ (3.4) satisfy the assumption

$$\sup_h \sum_{r,s=-N}^{N} |\kappa(h,r,s)| = O(N^{2d}). \tag{3.7}$$

In cases where condition (3.7) is difficult to check, a weaker Assumption L′ can be used, which nevertheless allows to prove the consistency of the tests.

Assumption L′**.** The stationary sequence $\{X_k\}$ satisfies (3.5) and (3.6).

3.1. Asymptotics for the modified R/S *statistic*

We begin with the asymptotic distribution of the modified R/S statistic (2.1) under the short memory hypothesis. In the sequel we shall denote by q the bandwidth parameter such that $q \to \infty$, $q/N \to 0$ as $N \to \infty$.

The asymptotics below follow using convergence (3.2) or (3.6), respectively, continuous mapping theorem and asymptotic properties of estimate $\hat{s}^2_{N,q}$ given in Theorem 3.1.

Proposition 3.1. *Suppose the sequence* $\{X_k\}$ *satisfies Assumption S. Then*

$$N^{-1/2}Q_N \;\Rightarrow\; U_{R/S}, \tag{3.8}$$

where

$$U_{R/S} = \max_{0\leqslant t\leqslant 1} W^0(t) - \min_{0\leqslant t\leqslant 1} W^0(t) \tag{3.9}$$

and $W^0(t) = W(t) - tW(1)$ *is a Brownian bridge.*

Remark 3.1. The distribution function of the random variable $U_{R/S}$ (3.9) (see Feller, 1951; Kuiper, 1960) is

$$F_{U_{R/S}}(x) = 1 + 2\sum_{k=1}^{\infty}(1-4k^2x^2)e^{-2k^2x^2}. \tag{3.10}$$

Critical values at any significance level can easily be obtained (see Table 2 in Lo, 1991) and it can be shown that

$$EU_{R/S} = \sqrt{\pi/2},\ \ \mathrm{Var}\, U_{R/S} = \pi(\pi-3)/6.$$

We note that distribution function (3.10) coincides with the limiting distribution in the Kuiper's (1960) goodness-of-fit test statistic.

Proposition 3.2 below shows that the test is consistent against long memory alternatives, and, in principle, can be used to study the power of the test.

Proposition 3.2. *Suppose that the sequence* $\{X_k\}$ *satisfies Assumption* L. *Then*

$$(q/N)^d N^{-1/2} Q_N \Rightarrow Z_{R/S}, \tag{3.11}$$

where

$$Z_{\mathrm{R/S}} = \sup_{0 \leqslant t \leqslant 1} W^0_{1/2+d}(t) - \inf_{0 \leqslant t \leqslant 1} W^0_{1/2+d}(t)$$

and $W^0_H(t) = W_H(t) - tW_H(1)$ *is the fractional Brownian bridge with parameter* H.

Under weaker Assumption L', *the consistency of the test still holds*:

$$N^{-1/2} Q_N \xrightarrow{P} \infty. \tag{3.12}$$

3.2. Asymptotics for the KPSS statistic

The asymptotic distribution of the statistic T_N defined by (2.4) can be handled in a similar way as that of the R/S statistic. We restrict ourselves to stating the results.

Proposition 3.3. *Under Assumption S*,

$$T_N \Rightarrow U_{\mathrm{KPSS}}, \tag{3.13}$$

where

$$U_{\mathrm{KPSS}} = \int_0^1 (W^0(t))^2 \,\mathrm{d}t.$$

Remark 3.2. The distribution function of the random variable U_{KPSS} has a series expansion in terms of special functions which converges very fast and can be used to tabulate the distribution of U_{KPSS}, see Kiefer (1959). The direct simulations of Kwiatkowski et al. (1992) which give slightly imprecise critical values, are unnecessary.

The random variable U_{KPSS} admits a series representation (see Rosenblatt, 1952)

$$U_{\mathrm{KPSS}} = \frac{1}{\pi^2} \sum_{j=1}^{\infty} \frac{Y_j^2}{j^2}, \tag{3.14}$$

where the Y_j are independent standard normal variables. Using (3.14), it is easy to verify that

$$\mathrm{E} U_{\mathrm{KPSS}} = 1/6, \quad \mathrm{Var}\, U_{\mathrm{KPSS}} = 1/45.$$

Proposition 3.4. *Under Assumption L*,

$$(q/N)^{2d} T_N \Rightarrow Z_{\mathrm{KPSS}}, \tag{3.15}$$

where

$$Z_{\mathrm{KPSS}} = \int_0^1 (W^0_{1/2+d}(t))^2 \,\mathrm{d}t.$$

Under Assumption L$'$,

$$T_N \xrightarrow{P} \infty. \tag{3.16}$$

3.3. Asymptotics for the V/S statistic

Proceeding as in Sections 3.1 and 3.2 we obtain the following results.

Proposition 3.5. *Under Assumption* S,

$$M_N \Rightarrow U_{\mathrm{V/S}}, \tag{3.17}$$

where

$$U_{\mathrm{V/S}} = \int_0^1 (W^0(t))^2 dt - \left(\int_0^1 W^0(t)\,\mathrm{d}t\right)^2.$$

Remark 3.3. The distribution function of the random variable $U_{\mathrm{V/S}}$ is given by the formula

$$F_{U_{\mathrm{V/S}}}(x) = 1 + 2\sum_{k=1}^{\infty}(-1)^k \mathrm{e}^{-2k^2\pi^2 x} \tag{3.18}$$

established by Watson (1961) in the context of goodness-of-fit tests on a circle. Note that $F_{U_{\mathrm{V/S}}}(x) = F_K(\pi\sqrt{x})$, where F_K is the asymptotic distribution function of the standard Kolmogorov statistic $\sup_{0<t<1}\sqrt{N}(\hat{F}_N(t) - t)$. Moreover, $U_{\mathrm{V/S}}$ admits the representation

$$U_{\mathrm{V/S}} = \frac{1}{4\pi^2}\sum_{j=1}^{\infty}\frac{(Y_{2j-1}^2 + Y_{2j}^2)}{j^2},$$

where the Y_j are independent standard normal variables, and so $U_{\mathrm{V/S}} = (U_1 + U_2)/4$, where U_1 and U_2 are two independent copies of the random variable U_{KPSS} appearing in (3.13). It follows that

$$\mathrm{E}U_{\mathrm{V/S}} = 1/12, \quad \mathrm{Var}\, U_{\mathrm{V/S}} = 1/360.$$

Thus the statistic M_N has a smaller variance than T_N.

Proposition 3.6. *Under Assumption L*,

$$(q/N)^{2d} M_N \Rightarrow Z_{\mathrm{V/S}}, \tag{3.19}$$

where

$$Z_{\mathrm{V/S}} = \int_0^1 (W^0_{1/2+d}(t))^2\,\mathrm{d}t - \left(\int_0^1 W^0_{1/2+d}(t)\,\mathrm{d}t\right)^2.$$

Under Assumption L′,

$$M_N \xrightarrow{P} \infty. \tag{3.20}$$

3.4. Asymptotic behavior of the variance estimator $\hat{s}^2_{N,q}$

The following result holds for the estimator $\hat{s}^2_{N,q}$ (2.2):

Theorem 3.1. *Let $\{X_k\}$ be a fourth order stationary process and let $q \to \infty$, $q/N \to 0$, as $N \to \infty$.*

(i) (*Short memory case*) *If (3.1) and (3.3) hold, then*

$$\hat{s}^2_{N,q} \xrightarrow{P} \sigma^2 := \sum_j \gamma_j. \tag{3.21}$$

(ii) (*Long memory case*) *If (3.5) and (3.7) hold, then*

$$q^{-2d} \hat{s}^2_{N,q} \xrightarrow{P} c_d^2 := \frac{c}{d(2d+1)}. \tag{3.22}$$

(*The constant c in (3.22) is the same as in (3.5).*)

(iii) *Under Assumption L',*

$$N^{-2d} \hat{s}^2_{N,q} \xrightarrow{P} 0. \tag{3.23}$$

Remark 3.4. Note that if $f(\lambda)$ denotes spectral density of $\{X_k\}$, then $\sigma^2 = 2\pi f(0)$, so $(2\pi)^{-1}\hat{s}^2_{N,q}$ is an estimator of $f(0)$. The assumptions of Theorem 3.1 (i) are weaker than the usual assumptions required in this context. For example, Theorem 9.3.4 of Anderson (1971) states, in particular, that if (3.1) and the condition

$$\sum_{h,r,s=-\infty}^{\infty} |\kappa(h,r,s)| < \infty \tag{3.24}$$

hold, then

$$\lim_{N\to\infty} \frac{N}{q} \operatorname{Var} \hat{f}_N(0) = 2 f^2(0) \int_{-1}^{1} K^2(x)\,dx, \tag{3.25}$$

provided $N \to \infty$, $q \to \infty$ and $q/N \to 0$. In (3.25) $\hat{f}_N(\lambda)$ is a kernel estimator of a spectral density $f(\lambda)$ defined by

$$\hat{f}_N(\lambda) = \frac{1}{2\pi} \sum_{|j|\leqslant q} K\left(\frac{j}{q+1}\right) \cos(\lambda j)\hat{\gamma}_j, \tag{3.26}$$

where the sample covariances $\hat{\gamma}_j$ are defined by (2.3), and $K(\cdot)$ is a continuous symmetric function on $[-1,1]$. Observe that setting $K(x) = 1 - |x|$ and $\lambda = 0$ in (3.26) we have $2\pi \hat{f}_N(0) = \hat{s}^2_{N,q}$, cf. (2.2). Condition (3.24) is stronger than (3.3) and more difficult to verify.

Remark 3.5. Examination of the proof of Theorem 3.1, see (6.6), shows that assumptions (3.3) and (3.7) can be replaced by more general condition

$$\sum_{r,s=-N}^{N} |\kappa(h,r,s)| \leqslant C N^{\alpha}$$

uniformly in $1 \leqslant h \leqslant N$ and N at the expense of a stronger assumption on $q \to \infty$: $q/N^{1-\alpha} \to 0$ $(0 \leqslant \alpha < 1)$ in the short memory case (i) and $q/N^{(1-\alpha)/(1-2d)} \to 0$ $(2d \leqslant \alpha < 1)$ in the long memory case (ii).

To illustrate the range of applicability of general Assumptions S and L we consider two examples. First of them concerns the well-known class of *linear sequences*

$$X_k = \sum_{j=-\infty}^{\infty} a_j \varepsilon_{k-j}, \tag{3.27}$$

where the a_j are real weights, $\sum_j a_j^2 < \infty$, and the ε_j are iid random variables with zero mean, unit variance and finite fourth moment $\mathrm{E}\varepsilon_0^4 < \infty$. For linear process (3.27) we have $\gamma_k = \sum_{j=-\infty}^{\infty} a_{j+k} a_j$ and

$$\kappa(h,r,s) = (\mathrm{E}\varepsilon_0^4 - 3) \sum_{k=-\infty}^{\infty} a_k a_{k+h} a_{k+r} a_{k+s}. \tag{3.28}$$

If the coefficients a_j satisfy condition

$$\sum_j |a_j| < \infty, \tag{3.29}$$

then $\{X_k\}$ satisfies short memory Assumption S.

Further, if

$$a_j \sim c j^{d-1} \quad (j \to \infty) \tag{3.30}$$

for some $c > 0$ and $0 < d < 1/2$, then $\{X_k\}$ satisfies long memory Assumption L. Indeed, in case when (3.30) holds, convergence (3.6) is well-known in the long memory literature, see, e.g. Davydov (1970). To verify (3.7), using (3.28) we get

$$k_N := \sum_{r,s=-N}^{N} |\kappa(h,r,s)| \leqslant C \sum_{|r|,|s| \leqslant N} \sum_{u=-\infty}^{\infty} |a_u a_{u+h} a_{u+r} a_{u+s}|$$

$$\leqslant C \left(\sum_{|u| \leqslant 2N} |a_u a_{u+h}| \sum_{|r|,|s| \leqslant 3N} |a_r||a_s| + \sum_{|u| > 2N} |a_u a_{u+h}| \sum_{|r|,|s| \leqslant N} |a_{u+r}||a_{u+s}| \right).$$

Since $\sum_u a_u^2 < \infty$, $\sum_{|u| \leqslant N} |a_u| \leqslant CN^d$ and $|a_{u+r}| \leqslant CN^{-1+d}$ for $|u| > 2N$, $|r| \leqslant N$, we get

$$k_N \leqslant C \left(\sum_{u=-\infty}^{\infty} a_u^2 \right) \left(\left(\sum_{|r| \leqslant 3N} |a_r| \right)^2 + N^{2d} \right) \leqslant CN^{2d}.$$

Besides that, under (3.30) the covariance $\gamma_k = \mathrm{Cov}(X_k, X_0)$ has property (3.5), i.e.

$$\gamma_k \sim C_d k^{2d-1}, \tag{3.31}$$

with $C_d = c^2 4^{-d} \pi^{-1/2} \Gamma(d) \Gamma(1/2 - d)$. The relationship between the constants c_d in (3.6) and C_d in (3.31) is $c_d^2 = C_d/(d(2d+1))$, and is easy to establish.

Observe, that relation (3.30), and hence also the long memory Assumption L, are satisfied by fractional ARIMA (p,d,q) sequence defined by

$$\Phi_p(L)(1-L)^d X_k = \Theta_q(L)\varepsilon_k \quad (0<d<1/2), \tag{3.32}$$

where $\Phi_p(L)$ and $\Theta_q(L)$ are polynomials of degree p and q, respectively, $\Phi_p(\cdot)$ satisfying the usual root requirement. In this case

$$a_j \sim \frac{\Theta_q(1)}{\Phi_p(1)\Gamma(d)}\, j^{d-1},$$

see Theorem 11.10 of Gourieroux and Monfort (1997).

The second example concerns the so-called *linear ARCH* (∞) or LARCH (∞) process, introduced by Robinson (1991).

As mentioned in the introduction, the classical ARCH(∞) model does not posses long memory, so in order to have an ARCH-type model nesting both short and long memory processes, a different specification is needed. The LARCH model is defined by

$$r_k = \sigma_k\varepsilon_k, \qquad \sigma_k = \alpha + \sum_{j=1}^{\infty}\beta_j r_{k-j}, \tag{3.33}$$

where $\{\varepsilon_k,\ k\in\mathbf{Z}\}$ is a sequence of zero mean finite variance iid random variables, α is a real number and the weights β_j determine the memory structure: if they decay sufficiently fast, we have a short memory model, if they decay hyperbolically, a long memory model investigated theoretically by GRS (2000) is obtained. A rigorous mathematical theory exists only for the long memory LARCH. We conjecture that if the β_j are absolutely summable (e.g. exponentially decaying), with $\sum|\beta_j|$ bounded by an appropriate constant and the innovations have moments of sufficiently high order, then the r_k^2 satisfy Assumption S. In simulations in Section 4, we consider the short memory LARCH model with the β_j given by

$$1+\sum_{j=1}^{\infty}\beta_j L^j = \frac{1-\theta L}{1-\phi L}, \ |\phi|<1. \tag{3.34}$$

The weights β_j for the long memory LARCH model satisfy

$$\beta_j \sim c j^{d-1}, \quad 0<d<1/2, \quad c\neq 0. \tag{3.35}$$

It has been shown that (essentially because of (3.35)) covariances of the squares r_k^2 decay hyperbolically and are not absolutely summable, and a functional limit theorem holds. More precisely, GRS (2000) proved that under assumption

$$L(\mathrm{E}\varepsilon_0^4)^{1/2}\sum_{j=1}^{\infty}\beta_j^2 < 1, \tag{3.36}$$

(where $L=7$ if the ε_k are Gaussian and $L=11$ in other cases), there is a stationary solution to Eqs. (3.33), (3.35) such that the squares $X_k=r_k^2$ satisfy

$$\mathrm{Cov}(X_k,X_0)\sim Ck^{2d-1}, \tag{3.37}$$

where C is a positive constant, and convergence (3.6) holds.

Hence, the squares r_k^2 of LARCH(∞) model (3.33), (3.35) possess two essential features of long memory: hyperbolically decaying non-summable covariances and attraction to the fractional Brownian motion, and satisfy Assumption L$'$. For brevity, we shall call the model $X_k = r_k^2$ (3.33), satisfying (3.35) and (3.36), LM LARCH(∞) model.

Although we expect that in LM LARCH(∞) case Assumption L holds as well, the verification of cumulant condition (3.7) remains an open technical problem.

In simulations in Section 4, we use for the LM LARCH(∞) model the coefficients β_j which satisfy

$$\beta_j = \beta_0 c_j, \quad 1 + \sum_{j=1}^{\infty} c_j L^j = (1-L)^{-d} \frac{1-\theta L}{1-\phi L}, \ |\phi| < 1. \tag{3.38}$$

The constant β_0 allows us to control the magnitude of the left-hand side of (3.36).

Estimation of the long memory parameter d in (3.33), (3.35) was discussed in Giraitis et al. (2000b).

4. Empirical size and power

We present in this section the results of a simulation study examining the finite sample performance of the test procedures developed in the previous sections. As mentioned in the Introduction, we see the main contribution of the present paper in advancing the relevant theory, so the empirical experiments presented below will leave some open questions which we hope will stimulate further research. Our goal is merely to form some idea on how the V/S test compares to the R/S and KPSS tests. Realizations of lengths $N = 500$ and $N = 1000$ are considered to make results comparable with those of Lo (1991) and Lee and Schmidt (1996).

For the modified R/S test the critical values can be found in Table 2 of Lo (1991), for the KPSS test in Tables 3 and 4 of Kiefer (1959), whereas those for the V/S test can be easily determined using a Kolmogorov statistic table, see Remark 3.3.

The tables below show the percentage of replications in which the rejection of a short memory null hypothesis was observed. Thus if a data generating process belongs to the null hypothesis, the tables show the empirical test sizes, and if it is a long memory process the tables show the empirical power of the tests. Several moving average and ARCH type models are considered. We first discuss the simulation results for linear and LARCH models, and finally we provide a simple application to exchange rate data.

For all simulations, the sequence of uniform deviates used for generating the error terms succeeds Marsaglia's (1996) DIEHARD tests.

I. *Linear sequences.* Table 1 presents the empirical sizes of the tests in AR(1) model $X_k = \phi X_{k-1} + \varepsilon_k$ and shows their dependence on the bandwidth parameter q. The ε_k are iid standard normal variables and ϕ takes values 0, 0.5, 0.8.

Table 1 shows that as parameter ϕ increases, what leads the increase in short range dependence, larger values of q must be used to obtain a correct size. Although in the iid case reasonable empirical size is obtained for $q = 0, \ldots, 10$, for large values of ϕ larger q should be used ($20 \leqslant q \leqslant 30$).

Table 1
Empirical test sizes (in %) of the modified R/S, KPSS, V/S and Lobato–Robinson statistics (where m_{opt} denotes the optimal automatic bandwidth) of the sequence X_k under the null hypothesis of AR(1) model, $X_k = \phi X_{k-1} + \varepsilon_k$, with standard normal innovations ε_k. Each row is based on 10 000 replications

		$\phi = 0$			$\phi = 0.5$			$\phi = 0.8$		
N	q	10%	5%	1%	10%	5%	1%	10%	5%	1%
Modified R/S statistic										
500	0	7.50	3.63	0.57	77.46	66.58	44.09	99.83	99.38	96.62
	1	7.28	3.42	0.55	45.98	33.64	15.43	93.88	88.73	71.18
	2	7.19	3.29	0.52	30.07	19.73	6.80	80.32	69.91	45.75
	5	6.88	2.94	0.34	14.06	7.46	1.46	43.41	30.87	11.60
	10	6.24	2.24	0.20	8.25	3.40	0.42	18.32	10.07	1.52
	20	5.12	1.38	0.06	4.83	1.39	0.05	6.32	2.01	0.03
	30	3.95	0.77	0.00	3.16	0.51	0.00	2.92	0.44	0.00
1000	0	7.96	3.92	0.71	81.21	70.90	48.68	99.96	99.74	98.19
	1	7.94	3.86	0.68	50.39	37.45	17.99	95.99	92.40	80.06
	2	7.92	3.85	0.63	33.67	22.47	8.37	85.23	76.61	55.61
	5	7.61	3.74	0.62	16.64	9.31	2.51	50.28	37.55	18.41
	10	7.19	3.50	0.47	10.62	5.03	0.94	24.54	15.19	4.49
	20	6.88	2.93	0.31	7.42	3.36	0.35	10.98	5.22	0.76
	30	6.34	2.42	0.19	6.06	2.42	0.18	7.19	2.95	0.24
KPSS statistic										
500	0	10.09	4.97	0.96	50.26	36.91	19.24	93.07	85.83	66.20
	1	10.00	4.99	0.96	31.73	21.51	8.60	73.88	60.61	37.74
	2	9.93	4.96	0.93	24.28	14.76	5.21	58.60	44.59	24.94
	5	9.71	5.00	0.83	15.99	8.90	2.40	35.46	24.11	9.97
	10	9.82	4.90	0.80	12.82	6.78	1.36	22.72	13.35	4.31
	20	9.71	4.56	0.54	11.27	5.38	0.85	15.04	8.12	1.61
	30	9.68	4.24	0.50	10.59	4.91	0.56	13.03	6.43	0.90
1000	0	10.17	5.03	1.18	51.12	38.21	19.67	93.28	85.48	67.30
	1	10.23	4.93	1.15	32.89	21.98	8.80	74.48	61.56	39.75
	2	10.18	4.93	1.10	24.88	15.30	5.18	59.63	45.97	25.92
	5	10.16	4.99	1.06	16.55	9.29	2.60	36.84	25.08	10.53
	10	10.31	4.85	1.00	13.13	6.73	1.70	23.42	14.20	4.60
	20	9.97	4.70	0.86	11.60	5.74	1.08	15.86	8.61	2.17
	30	9.98	4.64	0.71	10.86	5.21	0.81	13.52	6.94	1.48
V/S statistic										
500	0	10.09	4.88	0.99	68.62	55.67	33.09	99.44	98.08	91.14
	1	9.81	4.74	0.94	43.09	30.81	13.82	90.85	83.23	63.10
	2	9.59	4.70	0.89	31.33	20.48	6.97	77.95	65.73	42.63
	5	9.64	4.55	0.72	18.58	10.47	2.61	48.51	35.27	15.92
	10	9.29	4.24	0.54	13.59	6.64	1.08	28.51	17.56	5.00
	20	8.35	3.44	0.24	10.43	4.44	0.41	16.89	7.92	1.06
	30	7.46	2.56	0.08	8.94	3.07	0.13	12.62	4.86	0.24

Table 1 (*continued*)

N	q	$\phi=0$			$\phi=0.5$			$\phi=0.8$		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
1000	0	10.02	5.07	1.01	69.22	56.23	33.64	99.29	98.29	92.26
	1	9.90	4.94	0.96	43.60	31.14	13.84	91.83	84.02	64.35
	2	9.88	4.91	0.97	31.62	20.38	7.72	78.70	66.90	43.86
	5	9.84	5.05	0.90	18.85	11.06	3.04	49.22	36.01	17.09
	10	9.94	4.87	0.82	14.19	7.46	1.61	28.94	18.56	6.17
	20	9.64	4.49	0.54	11.83	5.73	0.91	17.73	10.06	2.16
	30	9.38	3.97	0.38	10.69	4.94	0.55	14.46	7.26	1.02
Lobato–Robinson statistic										
500	m_{opt}	7.00	3.32	0.34	8.57	4.69	1.24	4.94	2.23	0.44
1000	m_{opt}	7.85	3.52	0.60	10.05	5.62	1.47	6.87	3.55	0.87

If $\phi \geqslant 0.5$, the tests with $q=0,1,2,5$ have a strong bias towards rejecting the short memory null hypothesis. Although for these small q all three tests perform poorly, KPSS and V/S tests provide good approximations for large $20 \leqslant q \leqslant 30$. On the other hand, Table 2 shows that q should not be chosen too large, since under the long memory alternative, as q increases, these statistics have a strong bias towards accepting the null hypothesis, i.e. show “spurious” short memory (for more discussion see Teverovsky et al. (1999a, b).

Tables 1, 2 include also the empirical sizes and powers of the Lobato–Robinson test for long memory which uses the automatic bandwidth selection procedure proposed by Lobato and Robinson (1998) (see also Kirman and Teyssière, 2002). We also compute the Lobato–Robinson statistic on a grid of bandwidths as did Lobato and Savin (1998). To save space, we only report the results for the automatic bandwidth.

We can see that this test performs remarkably well. It should be noted that for the special case of the AR(1) model, Andrews (1991) proposed a formula for the optimal choice of q. Lo (1991) showed that Andrews’ formula works well for the modified R/S test in cases of $\phi=0$ (then optimal $q=0$) and $\phi=0.5$. We did not check how the Andrews’ formula works for the KPSS and V/S tests, but it can be expected that it will yield fairly good results. For example for $N=1000$ and $\phi=0.5$ it implies $q=14$ whereas for $N=1000$ and $\phi=0.8$, $q=31$. It can be gleaned from Table 1 that these values of q would most likely yield empirical sizes comparable to those of the Lobato–Robinson test.

Table 2 compares the power of the tests under three long memory alternatives. We consider the fractional ARIMA$(0,d,0)$ model defined in (3.32) with $d=0.2$, $d=0.3$ and $d=0.4$. Unlike Lo (1991), who studied also anti-persistent alternatives ($d<0$), we are concerned only with long memory alternatives ($d>0$), and thus consider one-sided tests, so the critical value for a size α modified R/S test is the $(1-\alpha)$th quantile of the distribution tabulated in Table 2 of Lo (1991).

The V/S statistic has always higher power than the KPSS statistic, for all values of q. For $q=0$ and 1, the modified R/S statistic is slightly more powerful than the

Table 2
Empirical power of the tests (in %) based on the modified R/S, KPSS, V/S and Lobato–Robinson (where m_{opt} denotes the "optimal" automatic bandwidth) statistics. The alternatives considered are FARIMA $(0,d,0)$ with $d=0.2$, $d=0.3$ and $d=0.4$. Innovations ε_k are standard normal. Each row is based on 10 000 replications

		$d=0.2$			$d=0.3$			$d=0.4$		
N	q	10%	5%	1%	10%	5%	1%	10%	5%	1%
Modified R/S statistic										
500	0	86.71	79.50	63.37	98.02	96.50	91.25	99.80	99.59	98.49
	1	77.45	68.20	48.47	94.13	90.43	79.94	98.81	97.66	93.95
	2	69.79	59.18	38.94	89.55	83.85	69.09	97.01	94.57	87.66
	5	53.13	41.53	22.46	75.66	65.33	45.28	88.31	82.28	66.64
	10	37.90	26.80	9.42	56.47	44.01	22.94	71.62	61.03	39.28
	20	21.92	10.66	1.10	32.38	19.40	2.88	43.93	29.05	6.31
	30	11.37	3.06	0.00	16.21	5.10	0.01	21.92	7.51	0.01
1000	0	93.81	89.99	78.94	99.72	99.32	97.41	100.00	99.98	99.92
	1	88.20	82.28	66.04	98.53	97.02	92.26	99.92	99.78	98.92
	2	83.34	75.23	56.67	96.54	93.98	86.61	99.58	99.05	96.83
	5	70.29	59.51	39.25	89.63	83.97	69.62	96.98	94.48	87.65
	10	55.77	43.97	24.63	77.41	67.94	48.37	89.81	84.12	69.71
	20	39.21	27.28	9.78	57.67	45.82	24.30	72.73	62.50	41.32
	30	29.02	17.54	3.95	44.31	30.62	10.14	57.63	45.15	20.91
KPSS statistic										
500	0	70.90	59.71	40.11	91.44	85.22	69.77	98.13	96.01	89.00
	1	62.77	51.13	31.73	84.65	75.36	57.86	94.82	90.31	78.92
	2	57.79	45.30	26.88	78.75	68.38	49.95	91.17	84.73	70.42
	5	47.95	35.57	18.93	66.78	55.09	36.43	81.01	72.04	53.49
	10	39.28	27.67	12.99	55.35	43.27	24.72	69.41	57.78	38.42
	20	31.01	20.38	7.08	43.44	31.46	14.34	55.04	42.58	23.95
	30	26.43	16.43	4.36	36.77	25.10	8.89	46.85	34.44	15.16
1000	0	79.49	69.31	50.52	95.92	92.38	82.26	99.51	98.89	95.82
	1	72.25	61.24	41.50	91.86	85.88	71.94	98.17	96.21	89.51
	2	67.12	55.84	35.89	87.68	80.92	64.41	96.56	93.33	83.97
	5	57.54	45.34	26.94	79.01	68.91	50.32	90.55	84.52	70.43
	10	48.59	36.60	19.89	68.49	56.56	38.06	82.52	73.81	56.02
	20	39.51	27.98	13.06	55.80	43.91	25.76	70.06	59.04	40.20
	30	34.27	23.33	9.48	48.36	36.71	18.99	61.60	49.94	30.93
V/S statistic										
500	0	83.91	75.38	57.56	97.05	94.74	86.86	99.59	99.13	97.03
	1	75.79	65.84	45.93	93.28	88.45	75.58	98.50	96.91	91.88
	2	69.90	58.86	38.44	89.23	82.51	66.54	96.69	94.06	85.74
	5	58.04	45.68	26.45	78.07	68.27	49.01	90.01	84.04	68.61
	10	46.82	34.86	16.05	65.24	53.74	32.60	78.74	69.40	49.65
	20	35.71	23.08	6.23	50.06	36.91	14.79	62.64	50.10	26.04
	30	28.97	15.85	1.90	40.07	26.09	5.06	51.66	36.86	10.45
1000	0	90.68	84.98	69.50	99.39	98.27	94.30	99.97	99.93	99.53
	1	85.05	76.60	58.62	97.52	94.90	87.84	99.79	99.45	97.57
	2	80.15	70.54	51.52	95.31	91.79	81.39	99.30	98.43	94.76

Table 2 (*continued*)

		$d=0.2$			$d=0.3$			$d=0.4$		
N	q	10%	5%	1%	10%	5%	1%	10%	5%	1%
	5	69.44	58.40	38.39	89.08	82.11	66.81	96.67	93.73	85.41
	10	58.84	46.98	26.77	79.03	69.80	50.88	90.67	85.17	70.63
	20	46.96	34.48	15.34	65.32	54.12	32.68	79.46	70.10	50.70
	30	39.76	27.23	9.69	56.10	43.75	21.64	70.31	58.65	36.98
Lobato–Robinson statistic										
500	m_{opt}	88.58	82.89	68.18	95.89	93.39	84.91	97.14	95.16	88.80
1000	m_{opt}	98.11	96.86	91.96	99.85	99.65	98.36	99.93	99.84	99.29

V/S, the V/S statistic having more power for $q>2$. When q increases, the power of the modified R/S statistic deteriorates so much that it becomes less powerful than the KPSS statistic. The relative changes of the power of the test as a function of q is illustrated by the size–power curves, advocated by Davidson and MacKinnon (1998) and given in Figs. 1 and 2. These curves plot the empirical distribution function of the P-values of the test under an alternative hypothesis against the empirical distribution of the P-values of the tests under a null hypothesis, i.e., the power of the test is adjusted to the correct size. The null hypothesis is chosen so that it is the closest to the alternative hypothesis according to the Kullback–Leibler criterion. The null and alternative hypotheses are plotted for a common value of q for the modified R/S, V/S and KPSS statistics, and a common bandwidth m for the Lobato–Robinson test.

When $q=0$, the curve for the modified R/S test is slightly above the KPSS. For large q, the size–power curve of the modified R/S test is below the curves of the other tests.

The limited simulations presented here indicate that for linear models the V/S test achieves a somewhat better balance of size and power than the R/S and KPSS tests; R/S tests suffer from size distortion and small power for large q, and the KPSS test has almost uniformly smaller power. The optimal range of the values of q to be used in practice depends on ϕ and N, and for the AR(1) model can be roughly determined by Andrews's formula. These simulations however do not show what happens for other short memory models, so further empirical work is required. For the models studied here, the Lobato–Robinson test distinguishes between short and long memory hypotheses with very high precision and should clearly be used in conjunction with the R/S-type tests. Note that, besides long memory testing, the KPSS and V/S tests might be applicable to unit root testing (see Kwiatkowski et al., 1992; Lee and Amsler, 1997; Shin and Schmidt, 1992).

II. *LARCH sequences*. Similarly as in the linear case, we first compare the empirical sizes of the three tests. We consider a short memory LARCH(1,0,1) model

$$\sigma_k^2=\left(\alpha+\sum_{j=1}^{\infty}\beta_j r_{k-j}\right)^2, \tag{4.1}$$

where the coefficients β_j are given by (3.34). The innovations ε_k are standard normal.

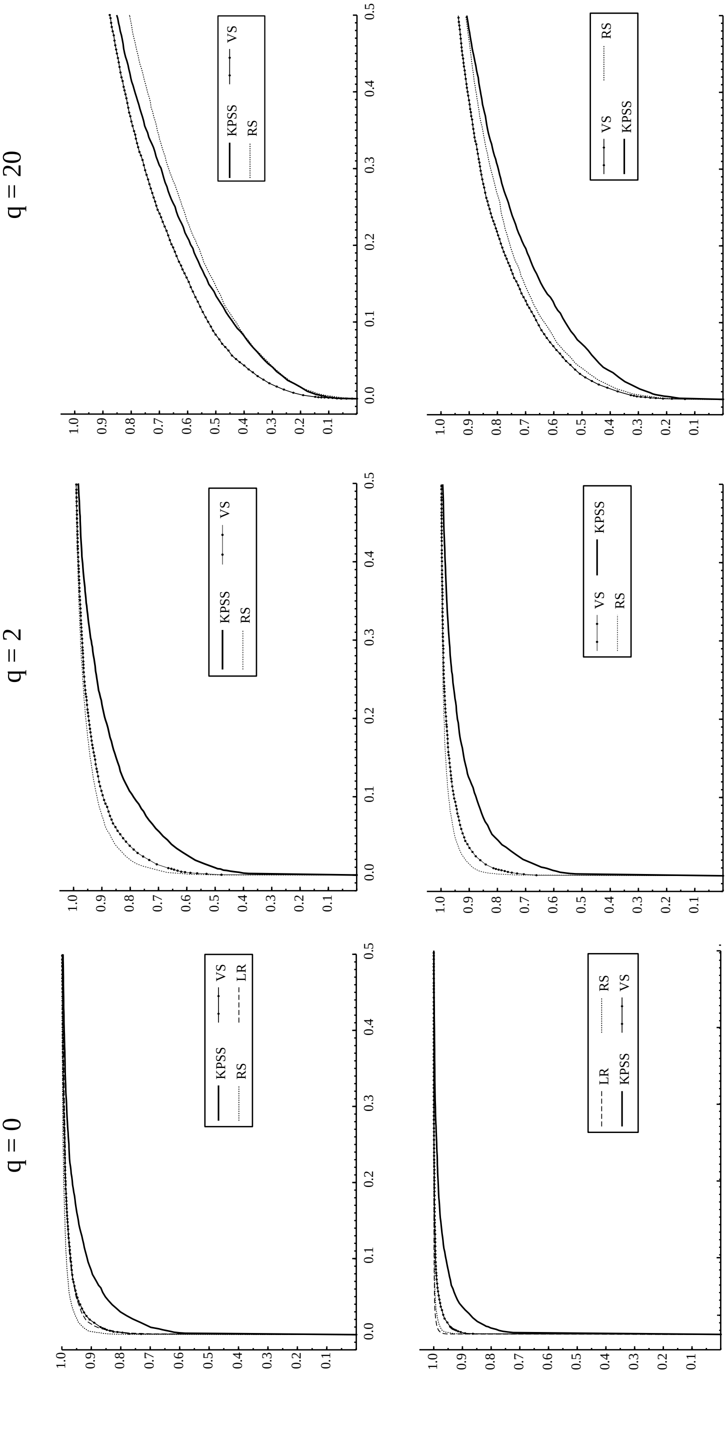

Fig. 1. Size–power curves of the tests for different values of q for iid $N(0,1)$–FARIMA$(0,d,0)$ models with $d = 0.3$. When $q = 0$, we also report the curve for the Lobato–Robinson statistic with the optimal bandwidth. First row corresponds to $N = 500$, second to $N = 100$.

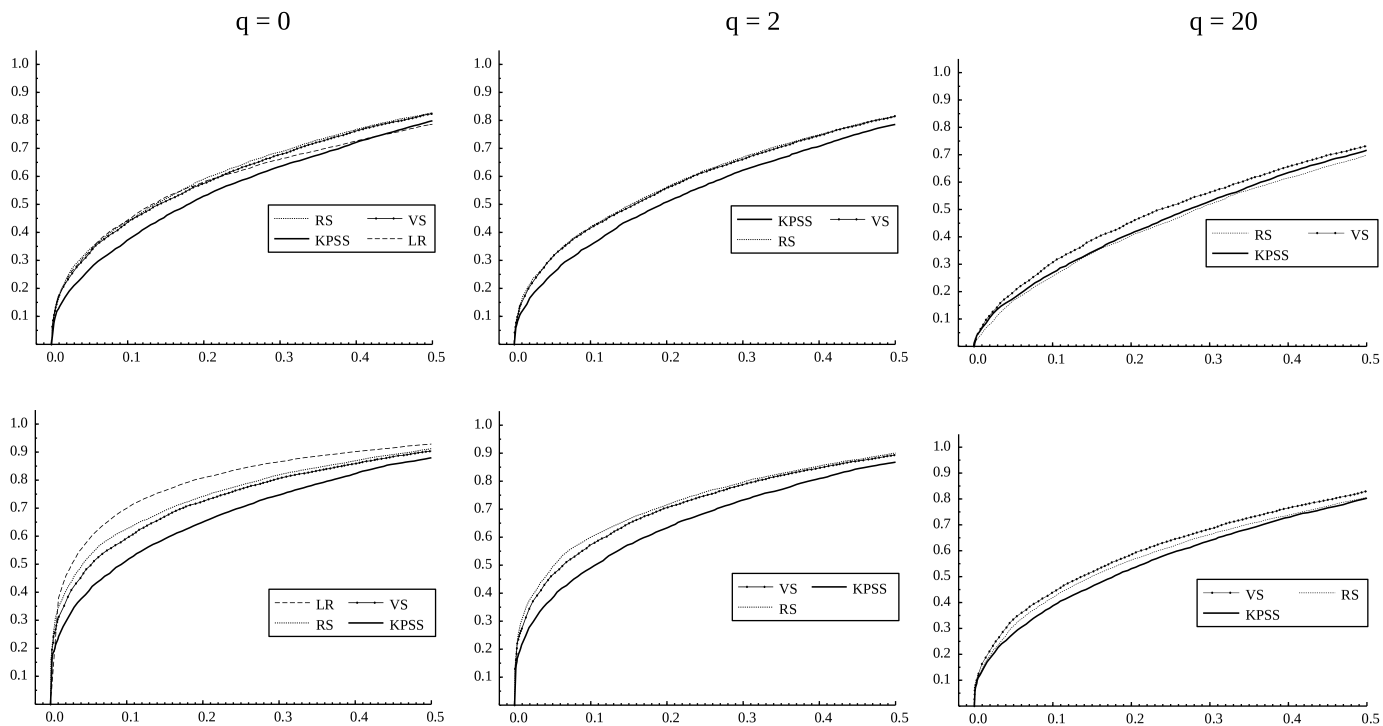

Fig. 2. Size–power curves of the tests for different values of q for LARCH(1,0,1)–LARCH($1,d,1$), models with $d=0.3$. When $q=0$, we also report the curve for the Lobato–Robinson statistic with the optimal bandwidth. First row corresponds to $N=500$, second to $N=1000$.

Table 3
Empirical test sizes (in %) of the modified R/S, KPSS and V/S statistics of the sequence r_k^2 under the null hypotheses of LARCH(1,0,1) model (4.1) with $\alpha = 0.1, \theta = 0.2$ and $\phi = 0.1$, cf. (3.38). The ε_k are standard normal. Each row is based on 10 000 replications

		Modified R/S statistic			KPSS statistic			V/S statistic		
N	q	10%	5%	1%	10%	5%	1%	10%	5%	1%
500	0	7.54	3.52	0.58	10.91	5.51	1.17	10.64	5.40	1.03
	1	6.83	3.00	0.48	10.63	5.31	1.04	10.18	4.99	0.89
	2	6.37	2.90	0.42	10.45	5.20	1.03	9.86	4.80	0.85
	5	5.97	2.47	0.30	10.28	5.02	0.94	9.53	4.55	0.75
	10	5.33	1.95	0.19	10.10	4.89	0.84	9.28	4.15	0.52
	20	4.20	1.26	0.04	10.01	4.75	0.66	8.51	3.33	0.23
	30	3.20	0.62	0.00	9.95	4.43	0.04	7.67	2.49	0.02
1000	0	8.26	3.94	0.60	10.28	5.05	1.05	10.79	5.26	1.09
	1	7.59	3.65	0.52	10.02	4.91	0.98	10.18	4.94	0.99
	2	7.14	3.41	0.50	9.90	4.77	0.92	9.84	4.78	0.96
	5	6.87	3.31	0.44	9.76	4.59	0.90	9.70	4.65	0.90
	10	6.57	3.05	0.35	9.64	4.61	0.83	9.62	4.38	0.73
	20	6.05	2.40	0.15	9.57	4.54	0.75	9.32	4.09	0.53
	30	5.53	1.95	0.08	9.59	4.40	0.56	9.07	3.72	0.37

Empirical sizes are reported in Table 3. Both V/S and KPSS tests have good size for the LARCH(1,0,1) null hypothesis, for $q = 0, \ldots, 10$ even when $N = 500$. There is a slight size distortion for the modified R/S test.

Tables 4 compares the power of the tests under the long memory alternatives defined in Section 3. The coefficients β_j in (3.33) are given by (3.38). Recall that for the Gaussian ε_k, the theory of testing under the alternative requires that

$$7(\mathrm{E}\varepsilon_0^4)^{1/2} \sum_{j=1}^{\infty} \beta_j^2 < 1 \tag{4.2}$$

(see (3.36)). Assumption (4.2) is only sufficient and the bound 1 is not the smallest possible. We set $\beta_0 = 1$ in (3.38). (If, for instance, β_0 is chosen so that the left-hand side of (4.2) is 0.99 then for $N = 500$ and significance level 10% the power of the test based on the V/S statistic is 25.25% if $d = 0.2$, 44.73% if $d = 0.3$ and 60.84% if $d = 0.4$).

According to the Kullback–Leibler criterion, the LARCH(1,0,1) is the process satisfying the null hypothesis which is the closest to the long memory LARCH$(1, d, 1)$ alternative. Size power curves for the LARCH(1,0,1)–LARCH$(1, d, 1)$ models are presented in Fig. 2.

The simulations indicate again that the V/S test may be an attractive alternative to the R/S and KPSS tests as its power is some 10–15% higher than that of the KPSS test and does not vary as rapidly with q as for the R/S test. Table 4 shows that in the LARCH setting the long memory can be fairly reliably detected for sample size $N = 1000$.

Table 4
Empirical power (in %) of the tests based on the modified R/S, the KPSS and the V/S statistics. The alternatives considered are the squares r_k^2 of the model (3.33), i.e., LARCH$(1,d,1)$ with the coefficients β_j of the form (3.38) with $\beta_0=1$, $\alpha=0.1$, $\theta=0.2$, $\phi=0.1$, $d=0.2$, $d=0.3$ and $d=0.4$ and standard normal innovations ε_k. Each row is based on 10 000 replications

		$d=0.2$			$d=0.3$			$d=0.4$		
N	q	10%	5%	1%	10%	5%	1%	10%	5%	1%
Modified R/S statistic										
500	0	72.10	62.55	43.57	79.23	71.73	54.72	74.44	65.70	48.29
	1	62.89	52.16	32.16	69.60	59.78	40.48	62.99	52.58	33.60
	2	55.97	44.12	24.71	61.76	50.67	30.40	53.45	42.72	23.80
	5	41.50	29.37	11.87	44.39	32.61	13.91	35.79	24.55	8.77
	10	27.18	16.53	3.57	27.64	16.25	3.75	20.06	10.66	2.09
	20	12.84	4.85	0.15	10.73	4.17	0.11	6.74	2.27	0.05
	30	5.56	0.93	0.00	4.07	0.61	0.00	2.26	0.32	0.00
1000	0	86.73	80.31	65.69	90.67	85.84	73.39	84.45	78.38	64.16
	1	79.78	72.20	54.25	84.21	77.21	61.78	76.12	67.61	50.41
	2	74.60	65.18	45.63	78.32	69.70	51.83	68.53	58.87	40.77
	5	61.89	50.60	29.86	64.31	53.41	33.22	52.47	40.96	22.24
	10	48.22	35.51	16.44	48.50	36.12	17.17	36.17	24.58	9.89
	20	31.43	19.20	5.87	29.32	17.80	5.09	19.06	10.99	2.28
	30	20.98	11.20	1.83	18.04	9.35	1.51	11.44	4.68	0.62
KPSS statistic										
500	0	64.49	53.42	33.95	74.59	64.10	45.04	73.97	62.98	43.14
	1	58.26	46.39	28.07	67.62	56.02	36.88	66.53	54.41	33.98
	2	54.11	41.98	23.75	62.88	50.55	31.33	60.85	48.16	28.07
	5	46.12	33.96	16.98	52.90	40.51	21.83	50.61	37.75	18.91
	10	39.01	26.88	11.59	44.21	31.97	14.12	42.29	28.90	11.71
	20	31.49	20.08	6.23	35.85	23.16	7.14	33.90	20.67	5.51
	30	27.69	15.98	3.49	31.11	17.96	3.82	29.19	16.02	2.56
1000	0	75.27	64.97	46.56	83.67	75.05	57.80	81.26	71.73	53.42
	1	69.48	58.08	39.83	77.77	67.63	48.96	74.72	64.08	43.91
	2	65.15	53.75	34.77	73.16	61.93	43.07	69.63	58.06	37.47
	5	56.88	44.98	26.73	63.88	51.91	31.99	59.74	46.79	26.90
	10	49.14	37.16	19.45	54.94	42.26	22.94	50.71	37.43	18.05
	20	40.48	28.62	12.81	44.76	32.00	14.47	41.28	28.03	10.67
	30	35.49	24.00	9.09	39.06	26.68	9.89	36.12	22.74	7.17
V/S statistic										
500	0	75.86	66.12	47.51	85.01	77.76	60.91	84.78	76.65	58.06
	1	69.57	58.72	39.25	79.06	69.82	49.73	77.82	67.43	46.44
	2	64.73	53.49	32.99	73.88	63.33	42.28	71.79	59.95	37.38
	5	55.34	43.17	21.38	63.25	50.52	26.96	59.49	45.27	21.78
	10	45.98	32.07	11.76	51.66	36.59	13.86	46.75	31.03	9.81
	20	33.90	19.46	3.28	36.78	21.31	3.22	31.46	15.81	1.93
	30	25.93	11.69	0.58	27.49	12.03	0.56	22.23	8.34	0.19

Table 4 (*continued*)

N	q	$d=0.2$ 10%	$d=0.2$ 5%	$d=0.2$ 1%	$d=0.3$ 10%	$d=0.3$ 5%	$d=0.3$ 1%	$d=0.4$ 10%	$d=0.4$ 5%	$d=0.4$ 1%
1000	0	86.45	79.60	64.03	92.96	87.83	74.90	90.98	84.69	70.43
	1	81.43	73.17	55.03	88.55	81.67	65.86	85.70	77.66	59.94
	2	77.49	68.28	48.78	84.55	76.86	58.71	81.19	71.56	51.68
	5	69.37	57.57	36.91	76.15	65.46	44.24	70.62	59.03	36.09
	10	59.79	47.25	25.75	66.23	52.88	30.35	59.97	45.87	21.77
	20	48.15	34.81	14.44	53.08	38.46	15.20	46.44	30.34	9.97
	30	40.87	26.85	8.27	44.26	28.71	8.15	37.49	21.29	4.58
Lobato–Robinson statistic										
500	m_{opt}	76.45	69.53	55.31	80.25	74.00	61.76	76.44	69.17	56.58
1000	m_{opt}	94.74	91.92	84.16	93.95	91.20	84.71	89.46	85.40	77.08

Interestingly, note from Table 4 that the power of all these tests decreases for large d, i.e., when the sufficient condition (4.2) is not satisfied, meaning that the series is apparently non-stationary or close to non-stationarity. This is in line with Vogelsang (1999), among others, who pointed out the behavior of the sample variance as a primary source of this non-monotonicity.

To simulate the series required to obtain Table 4 we used a pre-sample of 20000 observations which were recursively used for initiating the process. The fractional filter in (3.33) was truncated at the order 2000. Note that *Gaussian* linear fractionally integrated processes can be simulated exactly (without any truncation) by applying a form of a prediction error decomposition known also as the Durbin–Levinson algorithm, which was used in simulation of the FARIMA$(0,d,0)$ series.

III. *Application to exchange rate data.* We illustrate the theory and simulations by applying the tests to the squared returns r_k^2 of the exchange rate data. To make the results comparable with the simulations, we divide a series of four thousand daily returns on the Pound/Dollar exchange rate into four blocks of equal length as shown in Fig. 3. The corresponding P-values are displayed in Table 5. The evidence against the null hypothesis in favor of long memory alternative is strong for all blocks except perhaps the second one. The P-values, say, for the first block, are so small, that one is inclined to believe that the data exhibit some other forms of departure from the null than the alternative discussed in the present paper. This may be due to some form of non-stationarity, e.g. change in parameters of the conditional heteroskedastic process, or to probability tails of the marginal distributions of the r_k^2 which are heavier than those permitted by condition (4.2). In fact, the series of squared returns share the main second moments feature of long memory processes: a hyperbolic rate of decay of the autocorrelation function (and a pole in the periodogram near the zero frequency). The local Whittle estimator of the long memory parameter developed by Robinson (1995) yields $\hat{d}=0.3627$. However, Fig. 3 shows that such series do not display local spurious trends which are common in linear FARIMA-type models with such degree of persistence.

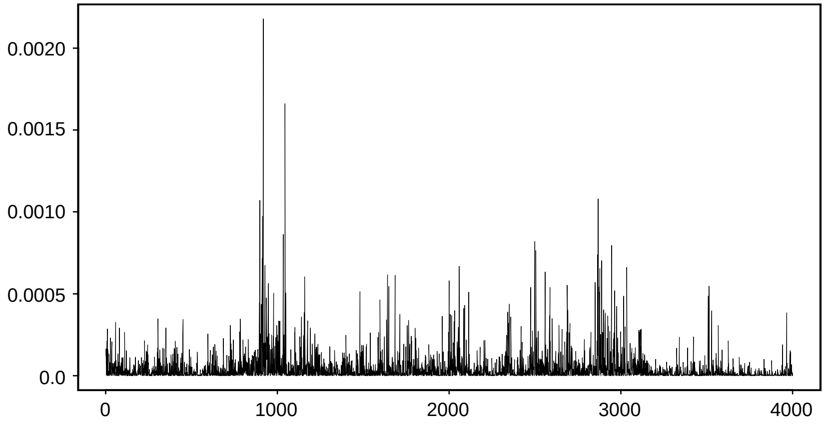

Fig. 3. Squares of returns on daily Pound/Dollar exchange rate. Four thousand observations ending 21 January 1997.

Table 5
P-values of the tests based on the modified R/S statistic, the KPSS statistic and the V/S statistics. For each value of q the four entries corresponding to a statistic give P-values for the four blocks shown in Fig. 3. For all blocks, except perhaps the second one, there is strong evidence against the null hypothesis

	Modified R/S	KPSS	V/S
$q=2$	0.000	0.000	0.000
	0.017	0.009	0.023
	0.001	0.003	0.001
	0.000	0.000	0.000
$q=5$	0.001	0.000	0.000
	0.074	0.024	0.067
	0.015	0.014	0.008
	0.001	0.000	0.001
$q=11$	0.017	0.002	0.001
	0.148	0.042	0.117
	0.055	0.032	0.024
	0.015	0.000	0.009

It is seen again that the R/S test is more sensitive to the choice of q and for large q appears to have a large probability of type II error. This accords with the findings of Teverovsky et al. (1999a, b) who looked at probabilities of type II error for linear sequences like fractional ARIMA and fractional Brownian motion. Note also that unlike in Table 4 the P-values for the V/S test are in most cases larger than for the KPSS test. This suggests again that the departure from the null hypothesis in the considered example cannot be exclusively explained by the long memory LARCH model considered in this paper.

5. Final comments

The goal of our paper was to introduce the new V/S test and to develop the theory of long memory testing based on the V/S and related statistics (KPSS and R/S). The results were derived for *general* classes of stationary processes satisfying Assumptions S or L. Special classes of processes which satisfy these assumptions (linear sequences and linear ARCH) were considered. In particular, Assumption S on the 4th order cumulants is well adjusted to the linear and LARCH processes and naturally satisfied without any additional restrictions besides the existence of corresponding 4th and 8th moments. Hence it may be expected that these tests are robust to the structure of the processes and applicable to a large class of sequences including mixing sequences. In special cases, the assumptions formulated in terms of cumulants might be easier to check than mixing assumptions, and assumptions of this type are widely used in time series analysis (see, e.g. Anderson, 1971; Brillinger, 1981). Even though mixing assumptions can be successfully used as well, it may be difficult to provide non-restrictive conditions for their validity in specific classes of processes. Note that to cover all possible memory scenarios we used the type of assumptions nesting both long and short memory sequences. Mixing is well adjusted to the weak dependence but it might not be suitable to characterize the long range dependence.

Although in special cases of linear processes and even ARCH processes the Lagrange multiplier type test of Lobato and Robinson (1998) based on the semiparametric estimation of the long memory parameter performs very well (see Lobato and Robinson, 1998; Kirman and Teyssière, 2002), its theoretical and empirical properties, especially the automatic bandwidth selection, requires additional investigation and, we believe that in a broad class of the null and alternative hypotheses, without assuming any particular structure of observations except stationarity, the R/S, KPSS, V/S-type tests are most useful.

The problem of choosing an optimal value of q is not yet solved. In the semiparametric estimation (linear models), a similar problem is the choice of a bandwidth m. The selection of the optimal q requires further empirical and theoretical work, but even without knowing the optimal value of q, if such exists at all, it might be useful to apply several tests and consider a range of values of q taking into account the length of the series.

Finally, the assumptions implying existence of the stationary solution to LARCH equations require relatively small β_j. Since these assumptions are not necessary, the case of "violated" theoretical condition (4.2) considered in simulations (Table 4) is also of interest.

Acknowledgements

We are grateful to Professor P. M. Robinson, the editor in charge, for constructive and helpful comments which lead to significant improvement of the paper. We also thank the referees.

Appendix A.

Proof of Theorem 3.1. The proofs in cases (i) and (ii) follow the same lines. Recall the definition of the sample covariances $\hat{\gamma}_j$ given in (2.3) and introduce the quantities

$$\tilde{\gamma}_j = \frac{1}{N} \sum_{i=1}^{N-|j|} (X_i - \mu)(X_{i+|j|} - \mu), \quad |j| < N.$$

Define also

$$Z_{k,l} := \sum_{i=k}^{l} (X_i - \mu).$$

Then

$$\hat{\gamma}_j - \tilde{\gamma}_j = \left(1 - \frac{|j|}{N}\right) (\bar{X}_N - \mu)^2 - N^{-1}(\bar{X}_N - \mu)(Z_{1,N-|j|} + Z_{|j|+1,N}) =: \delta_j.$$

Write $\hat{s}^2_{N,q} = v_{N,1} + v_{N,2}$, where

$$v_{N,1} = \sum_{|j| \leqslant q} \left(1 - \frac{|j|}{q+1}\right) \tilde{\gamma}_j, \quad v_{N,2} = \sum_{|j| \leqslant q} \left(1 - \frac{|j}{q+1}\right) \delta_j.$$

In case (i) set $d = 0$ and define $\sigma_*^2 = \sigma^2$; in case (ii) define $\sigma_*^2 = c_d^2$. It remains to show that

$$q^{-2d} v_{N,1} \xrightarrow{P} \sigma_*^2, \tag{A.1}$$

and

$$q^{-2d} v_{N,2} \xrightarrow{P} 0. \tag{A.2}$$

To obtain (A.2) notice that

$$\begin{aligned} \mathrm{E}|v_{N,2}| &\leqslant \sum_{|j| \leqslant q} \mathrm{E}|\delta_j| \\ &\leqslant \sum_{|j| \leqslant q} \mathrm{E}(\bar{X}_N - \mu)^2 + N^{-1}(\mathrm{Var}\,\bar{X}_N)^{1/2} \\ &\quad \sum_{|j| \leqslant q} ((\mathrm{E}Z^2_{1,N-|j|})^{1/2} + (\mathrm{E}Z^2_{|j|+1,N})^{1/2}). \end{aligned} \tag{A.3}$$

Since $\mathrm{E}Z^2_{k+1,l} = \sum_{i,i'=k+1}^{l} \gamma_{i-i'} \leqslant C\ (l-k)^{1+2d}$ and $\mathrm{Var}\,\bar{X}_N = N^{-2}\mathrm{E}Z^2_{1,N} \leqslant C\ N^{-1+2d}$ we obtain that (A.3) does not exceed

$$C \sum_{|j| \leqslant q} [N^{-1+2d} + N^{-1}N^{-1/2+d}(N - |j|)^{1/2+d}] \leqslant C(q/N)^{1-2d} q^{2d} = \mathrm{o}(q^{2d}).$$

To prove (A.1) note first that in case (i)

$$\mathrm{E}v_{N,1} = \sum_{|j| \leqslant q} \left(1 - \frac{|j|}{q+1}\right)\left(1 - \frac{|j|}{N}\right) \gamma_j \to \sum_{j=-\infty}^{\infty} \gamma_j \equiv \sigma^2 \quad (N \to \infty)$$

by (3.1). In case (ii)

$$\mathrm{E}v_{N,1} = \sum_{|j|\leqslant q}\left(1-\frac{|j|}{q+1}\right)\left(1-\frac{|j|}{N}\right)\gamma_j \sim \sum_{|j|\leqslant q}\left(1-\frac{|j|}{q+1}\right)\gamma_j$$

$$= (q+1)^{-1}\sum_{j,j'=1}^{q+1}\gamma_{j-j'} \sim q^{2d}C_d\int_0^1\int_0^1 |x-y|^{2d-1}\,\mathrm{d}x\,\mathrm{d}y = c_d^2 q^{2d}.$$

Thus, (A.1) follows if we show that

$$\mathrm{E}(v_{N,1}-\mathrm{E}v_{N,1})^2 = \mathrm{o}(q^{2d}).$$

Clearly

$$\mathrm{E}(v_{N,1}-\mathrm{E}v_{N,1})^2$$

$$= \sum_{|j|,|j'|\leqslant q}\left(1-\frac{|j|}{q+1}\right)\left(1-\frac{|j'|}{q+1}\right)\mathrm{Cov}(\tilde{\gamma}_j,\tilde{\gamma}_{j'})$$

$$\leqslant \sum_{|j|,|j'|\leqslant q}|\mathrm{Cov}(\tilde{\gamma}_j,\tilde{\gamma}_{j'})|. \tag{A.4}$$

Rewrite

$$\mathrm{Cov}(\tilde{\gamma}_j,\tilde{\gamma}_{j'}) = N^{-2}\sum_{i=1}^{N-|j|}\sum_{i'=1}^{N-|j'|}\mathrm{Cov}((X_i-\mu)(X_{i+|j|}-\mu),(X_{i'}-\mu)(X_{i'+|j'|}-\mu)).$$

Since

$$\mathrm{Cov}((X_i-\mu)(X_{i+|j|}-\mu),(X_{i'}-\mu)(X_{i'+|j'|}-\mu))$$

$$= \mathrm{Cum}(X_i,X_{i+|j|},X_{i'},X_{i'+|j'|}) + \gamma_{i-i'}\gamma_{i-i'+|j|-|j'|} + \gamma_{i-i'-|j'|}\gamma_{i-i'+|j|}, \tag{A.5}$$

we obtain from (A.4) and (A.5)

$$\mathrm{E}(v_{N,1}-\mathrm{E}v_{N,1})^2$$

$$\leqslant N^{-2}\sum_{|j|,|j'|\leqslant q}\sum_{i=1}^{N}\sum_{i'=1}^{N}|\mathrm{Cum}(X_i,X_{i+|j|},X_{i'},X_{i'+|j'|})|$$

$$+N^{-2}\sum_{|j|,|j'|\leqslant q}\sum_{i=1}^{N}\sum_{i'=1}^{N}(|\gamma_{i-i'}\gamma_{i-i'+|j|-|j'|}|$$

$$+|\gamma_{i-i'-|j'|}\gamma_{i-i'+|j|}|) =: i_{N,1}+i_{N,2}.$$

By (3.1) in case (i) and by (3.5) in case (ii) it follows that $\sum_{|i|\leqslant 2N}|\gamma_i| \leqslant CN^{2d}$. (Recall that in short memory case $d=0$.) Therefore

$$i_{N,2} \leqslant 2N^{-2}\sum_{|j|\leqslant q}\sum_{i=1}^{N}\sum_{|i'|,|j'|\leqslant 2N}|\gamma_{i'}|\,|\gamma_{j'}| \leqslant C(q/N)^{1-2d}q^{2d} = \mathrm{o}(q^{2d}).$$

For the term $i_{N,1}$, using the property of cumulants

$$\mathrm{Cum}(X_0, X_{|j|}, X_{i'-i}, X_{i'-i+|j'|}) = \mathrm{Cum}(X_0, X_{|j|}, X_{i'-i}, X_{i'-i+|j'|}),$$

we get

$$\begin{aligned} i_{N,1} &= N^{-2} \sum_{|j| \leqslant q} \sum_{i'=1}^{N} \left(\sum_{i=1}^{N} \sum_{|j'| \leqslant q} |\mathrm{Cum}(X_0, X_{|j|}, X_{i'-i}, X_{i'-i+|j'|})| \right) \\ &\leqslant N^{-1} \sum_{|j| \leqslant q} \left(\sum_{i,j'=-2N}^{2N} |\mathrm{Cum}(X_0, X_{|j|}, X_i, X_{j'})| \right) \\ &\leqslant C(q/N)^{1-2d} q^{2d} = \mathrm{o}(q^{2d}), \end{aligned} \tag{A.6}$$

by assumptions (3.3) and (3.7).

We now turn to the proof of part (iii). Since $\hat{s}^2_{N,q}$ is non-negative and $q/N \to 0$, it is enough to verify that $\mathrm{E}\hat{s}^2_{N,q} \leqslant Cq^{2d}$. We have

$$E\hat{s}^2_{N,q} = \sum_{|j| \leqslant q} \left(1 - \frac{|j|}{q+1}\right) \mathrm{E}\hat{\gamma}_j \leqslant \sum_{|j| \leqslant q} |\mathrm{E}\hat{\gamma}_j - \gamma_j| + \sum_{|j| \leqslant q} |\gamma_j|.$$

Using relation (3.5), it is not difficult to verify that (see, for example, relation (8) in Hosking (1996) or proof of relation (A.2))

$$|\mathrm{E}\hat{\gamma}_j - \gamma_j| \leqslant CN^{2d-1}, \tag{A.7}$$

where $C > 0$. Thus, by (A.7) and (3.5)

$$\mathrm{E}\hat{s}^2_{N,q} \leqslant C_1 \sum_{|j| \leqslant q} N^{2d-1} + C_2 \sum_{|j| \leqslant q} j^{2d-1} \leqslant C_3 q^{2d},$$

where we used the assumption $0 < 2d < 1$. □

References

Anderson, T.W., 1971. The Statistical Analysis of Time Series. Wiley, New York.

Andrews, D.W.K., 1991. Heteroscedasticity and autocorrelation consistent covariance estimation. Econometrica 59, 817–858.

Baillie, R.T., Bollerslev, T., Mikkelsen, H.O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3–30.

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

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

Brillinger, D.R., 1981. Time Series: Data Analysis and Theory. Holt, Rinehart and Winston, New York.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton.

Cheung, Y.W., 1993a. Long memory in foreign-exchange rates. Journal of Business and Economic Statistics 11, 93–101.

Cheung, Y.W., 1993b. Tests for fractional integration: A Monte Carlo investigation. Journal of Time Series Analysis 14, 331–345.

Cheung, Y.W., Lai, K.S., 1993. Do gold market returns have long memory? The Financial Review 28, 181–202.

Crato, N., de Lima, P.J.F., 1994. Long-range dependence in the conditional variance of stock returns. Economics Letters 45, 281–285.

Davidson, J., 2002. Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. Journal of Econometrics 106, 243–269.

Davidson, J., de Jong, R., 2000. Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. Econometrica 68, 407–423.

Davidson, R., MacKinnon, J.G., 1998. Graphical methods for investigating the size and power of hypothesis tests. The Manchester School 66, 1–26.

Davydov, Y.A., 1970. The invariance principle for stationary processes. Theory of Probability and its Applications 15, 487–498.

Ding, Z., Granger, C.W.J., 1996. Modeling volatility persistence of speculative returns: a new approach. Journal of Econometrics 73, 185–215.

Durbin, J., 1973. Distribution Theory for Tests Based on the Sample Distribution Function. Society for Industrial and Applied Mathematics, Philadelphia.

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

Geweke, J., Porter-Hudak, S., 1983. The estimation and application of long memory time models. Journal of Time Series Analysis 4, 221–238.

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

Giraitis, L., Kokoszka, P., Leipus, R., Teyssière, G., 2000b. Semiparametric estimation of the intensity of long memory in conditional heteroskedasticity. Statistical Inference for Stochastic Processes 3, 113–128.

Giraitis, L., Robinson, P., Surgailis, D., 2000c. A model for long memory conditional heteroskedasticity. Annals of Applied Probability 10, 1002–1024.

Giraitis, L., Kokoszka, P., Leipus, R., 2001. Testing for long memory in the presence of a general trend. Journal of Applied Probability 38, 1033–1054.

Goetzmann, W.N., 1993. Patterns in three centuries of stock market prices. Journal of Business 66, 249–270.

Gourieroux, C., Monfort, A., 1997. Time Series and Dynamic Models. Cambridge University Press, Cambridge.

Hauser, M.A., 1997. Semiparametric and nonparametric testing for long memory: a Monte Carlo study. Empirical Economics 22, 247–271.

Hosking, J.R.M., 1996. Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. Journal of Econometrics 73, 261–284.

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

Kiefer, J., 1959. K-sample analogues of the Kolmogorov–Smirnov and Cramér-v. Mises tests. Annals of Mathematical Statistics 30, 420–447.

Kirman, A., Teyssière, G., 2002. Microeconomic models for long-memory in the volatility of financial time series. Studies in Nonlinear Dynamics and Econometrics 5, 281–302.

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

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 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.

Lee, H.S., Amsler, C., 1997. Consistency of the KPSS unit root test against fractionally integrated alternative. Economics Letters 55, 151–160.

Lee, D., Schmidt, P., 1996. On the power of the KPSS test of stationarity against fractionally-integrated alternatives. Journal of Econometrics 73, 285–302.

Liu, Y.A., Pan, M.S., Hsueh, L.P., 1993. A modified R/S analysis of long-term dependence in currency futures prices. Journal of International Financial Markets, Institutions and Money 3, 97–113.

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

Lobato, I., Robinson, P.M., 1998. A nonparametric test for I(0). Review of Economic Studies 68, 475–495.
Lobato, I., Savin, N.E., 1998. Real and spurious long-memory properties of stock-market data (with comments). Journal of Business & Economic Statistics 16, 261–283.
Mandelbrot, B.B., 1972. Statistical methodology for non-periodic cycles: from the covariance to R/S analysis. Annals of Economic and Social Measurement 1, 259–290.
Mandelbrot, B.B., 1975. Limit theorems of the self-normalized range for weakly and strongly dependent processes. Zeitschrift für Wahrschein lichkeitstheorie und Verwandte Gebiete 31, 271–285.
Mandelbrot, B.B., Taqqu, M.S., 1979. Robust R/S analysis of long run serial correlation. 42nd Session of the International Statistical Institute, Manila, Book 2, pp. 69–99.
Mandelbrot, B.B., Wallis, J.M., 1969. Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. Water Resources Research 5, 967–988.
Marsaglia, G., 1996. DIEHARD: a battery of tests of randomness. http://stat.fsu.edu/pub/diehard.
Mikosch, T., Stărică, C., 1999. Change of structure in financial time series, long range dependence and the GARCH model. Preprint.
Robinson, P.M., 1991. Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. Journal of Econometrics 47, 67–84.
Robinson, P.M., 1995. Gaussian semiparametric estimation of long range dependence. The Annals of Statistics 23, 1630–1661.
Robinson, P.M., Henry, M., 1999. Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels. Econometric Theory 15, 299–336.
Robinson, P.M., Zaffaroni, P., 1998. Nonlinear time series with long memory: a model for stochastic volatility. Journal of Statistical Planning and Inference 68, 359–371.
Rosenblatt, M., 1952. Limit theorems associated with variants of the von Mises statistic. Annals of Mathematical Statistics 23, 617–623.
Shin, Y., Schmidt, P., 1992. The KPSS stationary test as a unit root test. Economics Letters 38, 387–505.
Teverovsky, V., Taqqu, M.S., Willinger, W., 1999a. A critical look at Lo's modified R/S statisitic. Journal of Statistical Planning and Inference 80, 211–227.
Teverovsky, V., Taqqu, M.S., Willinger, W., 1999b. Stock market prices and long-range dependence. Finance & Stochastics 3, 1–13.
Teyssière, G., 2002. Interaction models for common long-range dependence in asset prices volatilities. Long Range Dependent Stochastic Processes: Theory and Applications. Springer, Berlin, forthcoming.
Tsay, V.-S., 1998. On the power of Durbin–Watson statistic against fractionally integrated processes. Econometric Reviews 17, 361–386.
Vogelsang, T.J., 1999. Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series. Journal of Econometrics 88, 283–299.
Watson, G.S., 1961. Goodness-of-fit tests on a circle. Biometrika 48, 109–114.