Testing for long memory in the presence of a general trend *

Liudas Giraitis[†]
London School of Economics

Piotr Kokoszka
The University of Liverpool

Remigijus Leipus[‡]
Vilnius University,
Institute of Mathematics and Informatics

June 20, 2001

Abstract

The paper studies the impact of a broadly understood trend, which includes a change-point in mean and monotonic trends studied by Bhattacharya *et al.* (1983), on the asymptotic behaviour of a class of tests designed to detect long memory in a stationary sequence. Our results pertain to a family of tests which are similar to Lo's (1991) modified R/S test. We show that both long memory and nonstationarity (presence of trend or change points) can lead to rejection of the null hypothesis of short memory, so that further testing is needed to discriminate between long memory and and some forms of nonstationarity. We provide quantitative discription of trends which do and do not fool the R/S type long memory tests. We show, in particular, that a shift in mean of a magnitude larger than $N^{-1/2}$, where N is the sample size, affects the asymptotic size of the tests, whereas smaller shifts do not do so.

Key words: Long memory, trend, change-point.

Running head: Long memory and trend.

Corresponding Author: Liudas Giraitis, Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom; l.giraitis@lse.ac.uk

*Research partially supported by an EPSRC grant GR/L/78222 at the University of Liverpool

[†]Supported by the ESRC grant R000238212.

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

1 Introduction

Starting with the seminal work of Hurst (1951), which was subsequently refined by Mandelbrot and Wallis (1969), Mandelbrot (1972, 1975) and Mandelbrot and Taqqu (1979), among others, a continued interest in the so-called Hurst effect lasts up to this day. Following Mandelbrot and Taqqu (1979) and Bhattacharya *et al.* (1983), this effect can be defined by means of the R/S statistic as follows. Let $X_1, \ldots, X_N$ be a sequence of random variables. The R/S statistic Q_N is defined as $R_N/\hat{s}_N$, where R_N is the (adjusted) range,

$$R_N = \max_{1\le k\le N} \sum_{j=1}^{k} (X_j - \bar{X}_N) - \min_{1\le k\le N} \sum_{j=1}^{k} (X_j - \bar{X}_N), \tag{1.1}$$

and

$$\hat{s}_N^2 = \frac{1}{N} \sum_{j=1}^{N} (X_j - \bar{X}_N)^2 \tag{1.2}$$

is the sample variance. In (1.1), (1.2), $\bar{X}_N$ is the sample mean $N^{-1}\sum_1^N X_j$. Then by the *Hurst effect* we mean that for some $H > 0.5$, $N^{-H}Q_N$ converges in distribution as $N \to \infty$ to a random variable which is possibly degenerate but almost surely not 0. As has been well documented since the 1950's, for a large number of geophysical and hydrological records the plots of $\log Q_N$ against $\log N$ indicate that Q_N is asymptotically proportional to N^H with $H > 0.5$. The same phenomenon has recently been observed in telecommunication and computer networks, see e.g. Willinger *et al.* (1995) and references therein. It is also believed to be present in some financial and economic series, see e.g. Section 2.6 of Campbell *et al.* (1997).

One possible explanation for this effect is that the sample $X_1, \ldots, X_N$ is a part realization of a long-range dependent (or long memory) sequence, for example of a Gaussian stationary sequence with non-summable covariances. Although the stationarity of a sequence, i.e. the assumption that the data possess constant mean, variance and shift-invariant covariance function, may appear to be an over-idealization of real data, it often provides a reasonable and useful approximation to a possibly non-stationary data generating model. An alternative explanation of the Hurst effect is that the data have short memory but are perturbed by a slowly declining deterministic trend or abrupt changes in parameters, like e.g. changes in mean. To our knowledge the first comprehensive mathematical study of the Hurst effect for short memory models with trend was done by Bhattacharya *et al.* (1983) who assumed that the observations $X_1, X_2, \ldots$ are of the form

$$X_k = f(k) + Y_k, \tag{1.3}$$

where $\{Y_k\}$ is an ergodic sequence with summable covariances whose normalised partial sums converge to a Brownian motion (see (2.4) below). The sequence $\{f(k)\}$ is a deterministic trend. Their main result states that the Hurst effect for the short memory models with trend occurs if and only if

$$\Delta^{(N)} \sim cN^{H-1/2}, \quad H > 1/2, \quad c > 0,$$

where

$$\Delta^{(N)} = N^{-1/2}\Big[\max_{1\le k\le N} \sum_{j=1}^{k} (f(j) - \bar{f}_N) - \min_{1\le k\le N} \sum_{j=1}^{k} (f(j) - \bar{f}_N)\Big]$$

with $\bar{f}_N = N^{-1}\sum_1^N f(j)$, and $\sim$ indicates that the ratio of left and right-hand sides tends to 1 as $N \to \infty$. In particular, Bhattacharya *et al.* (1983) considered a monotonic trend of the form $f(k) = c'(m+k)^{\beta}$, where m is an arbitrary non-negative parameter and $c' \neq 0$. In this case the Hurst effect is present, unless $\beta \leq -1/2$ or $\beta = 0$. In the most subtle case $-1/2 < \beta < 0$, one obtains a model with trend decaying so slowly to zero that it can be confused with a long memory model. Assuming Gaussianity, Künsch (1986) suggested a procedure to discriminate between the monotonic trend model $X_k = f(k) + Y_k$, $f(k) \to 0$ and long-range dependence. The procedure is based on the periodogram which, as shown by Künsch (1986), behaves differently for the model (1.3) and a stationary long-range dependent sequence. In a similar setting, Heyde and Dai (1996) proposed a somewhat more informative smoothed periodogram approach and argued that periodogram based procedures are less sensitive to departures from stationarity than those based on R/S and its variants.

A situation where an observed series exhibits the Hurst effect, but a possible explanation for such a behaviour is believed to be different than long-range dependence, is often referred to as "spurious" long memory. For recent investigations of various aspects of the "spurious" long memory we refer to Teverovsky and Taqqu (1997), Lobato and Savin (1998), Bos *et al.* (1999), Mikosch and Stărică (1999), Granger and Hyung (1999), Diebold and Inoue (1999). As noted by Diebold and Inoue (1999), in certain circumstances some trends (e.g. structural change) and long memory are "effectively different labels for the same phenomenon, in which case attempts to label one as "true" and the other as "spurious" are of dubious value".

Similar situation also arises in the unit roots tests where structural breaks can be detected as a unit root and vice versa, see Perron (1990), Perron and Vogelsang (1992), Perron (1997) and Baum, Barkoulas and Caglayan (1999).

This paper focuses on the impact of a broadly understood trend on a class of statistical tests designed to verify the short memory null hypothesis against the long memory alternative. We use here a "classical" definition of long memory: roughly speaking a stochastic sequence is said to have long memory if it is covariance stationary and the autocovariance function is not absolutely summable. For interesting extensions of the concept of long memory we refer to Hall (1997) and Heyde and Yang (1997). By a short memory sequence we mean a covariance stationary sequence with absolutely summable autocovariance function. These rough definitions are made precise in the following sections. Unlike many earlier studies which focused on simulational evidence, we provide rigorous mathematical formulations. Some specific questions we seek to answer are: if the data come from a short memory model perturbed by a trend, how "large" must this trend be to lead to the detection of "spurious" long memory? What kind of trends may fool the long memory test? We provide a general answer in terms of the L^2 norm of the trend. Similarly, if there is a shift in the mean of a stationary short memory sequence, what must the magnitude of that shift be to have an impact on the test. It will be shown that a change of a constant magnitude which does not decrease with the sample size, will be regitered as long memory with probability approaching 1, as the sample size, N, tends to infinity. We show that a change of the magnitude $N^{-1/2}$ is on the border line: smaller changes do not (asymptotically) change the significance level. Our approach encompasses traditional linear models as well as a family of ARCH models for which conditions for the applicability of the R/S type tests are naturally satisfied.

Our results pertain to tests which are similar in spirit to Lo's (1991) test based on the *modified* R/S statistic $Q_N(q)$ defined as

$$Q_N(q) = \frac{R_N}{\hat{s}_{N,q}},$$

with R_N given in (1.1) and

$$\hat{s}^2_{N,q} \equiv (q_N + 1)^{-1} \sum_{i,j=1}^{q_N+1} \hat{\gamma}_{i-j} = \frac{1}{N} \sum_{j=1}^{N} (X_j - \bar{X}_N)^2 + 2 \sum_{j=1}^{q_N} \omega_j(q_N) \hat{\gamma}_j \tag{1.4}$$

with $\omega_j(q_N) = 1 - j/(q_N + 1)$ being the Bartlett weights and the $\hat{\gamma}_j$ being the sample covariances, $\hat{\gamma}_j = N^{-1} \sum_{i=1}^{N-j} (X_i - \bar{X}_N)(X_{i+j} - \bar{X}_N), \quad 0 \le j < N$. Note that the classical R/S statistic corresponds to $q_N = 0$. The second term in (1.4) was introduced by Lo (1991) in order to take into account short range dependence. The classical R/S analysis focused on estimating the Hurst coefficient as the limit of the ratio $\log Q_N(0)/\log N$, whereas Lo's (1991) test is concerned with a complex null hypothesis encompassing many forms of short memory including ARMA and other mixing sequences with mixing coefficients which decay sufficiently fast.

The present paper focuses mainly on the V/S statistic introduced by Giraitis, Kokoszka and Leipus (2000a) which has more desirable asymptotic and finite-sample properties than the modified R/S statistic. Our results can, however, be easily formulated in terms of the modified R/S statistic and other related statistics, like the KPSS statistic of Kwiatkowski *et al.* (1992). The general conclusions are the same for all statistics in this class. The V/S statistic we analyse is defined as $V_N/\hat{s}^2_{N,q}$, where

$$V_N = \frac{1}{N^2} \left[\sum_{k=1}^{N} \left(\sum_{j=1}^{k} (X_j - \bar{X}_N) \right)^2 - \frac{1}{N} \left(\sum_{k=1}^{N} \sum_{j=1}^{k} (X_j - \bar{X}_N) \right)^2 \right]. \tag{1.5}$$

Comparing to R/S statistic, the numerator (1.5) of the V/S statistic is based on the sample *variance* of the sums $S^*_{X,k} = \sum_1^k (X_j - \bar{X}_N)$, rather than their *range* (1.1).

In Section 2, we study the effect of a trend on the size of the tests. Section 3 focuses on the power of the tests. Important special cases of a trend and a change in mean are analysed in detail in Section 4. Proofs are collected in Section 6.

2 Short memory and trend

In this section we consider the following model

$$X_k = f^{(N)}(k) + Y_k, \quad k = 1, 2, \ldots, \tag{2.1}$$

where $\{Y_k\}$ is a short memory process and $\{f^{(N)}(k)\}$ is a deterministic trend. We now state the precise conditions we impose on the process $\{Y_k\}$ (Assumption A) and on the sequence $\{f^{(N)}(k)\}$ (Assumptions B and C) together with several examples.

Recall that the fourth-order cumulant of the random variables $\xi_1, \xi_2, \xi_3, \xi_4$ is defined by

$$\mathrm{Cum}(\xi_1, \xi_2, \xi_3, \xi_4) = E\left[\prod_{j=1}^{4} (\xi_j - E\xi_j) \right] - \mathrm{Cov}(\xi_1, \xi_2)\mathrm{Cov}(\xi_3, \xi_4) \tag{2.2}$$

$$-\mathrm{Cov}(\xi_1,\xi_3)\mathrm{Cov}(\xi_2,\xi_4)-\mathrm{Cov}(\xi_1,\xi_4)\mathrm{Cov}(\xi_2,\xi_3).$$

Throughout the paper, for $t \in [0,1]$, $W(t)$ stands for the standard Brownian motion and $W^0(t) = W(t) - tW(1)$ for the Brownian bridge. By $\xrightarrow{d}$ we denote the weak convergence of random variables, by $\xrightarrow{fdd}$ the weak convergence of finite dimensional distributions and by $\xrightarrow{D[0,1]}$ the weak convergence in the Skorokhod space $D[0,1]$.

ASSUMPTION A. (Short memory) $\{Y_k\}$ is a fourth-order stationary zero mean sequence satisfying conditions $(i)-(iii)$ below:

(i) the covariance function is absolutely summable:

$$\sum_{k=-\infty}^{\infty} |\mathrm{Cov}(Y_k,Y_0)| < \infty; \tag{2.3}$$

(ii) the functional CLT holds, i.e.

$$N^{-1/2}\sum_{j=1}^{[Nt]} Y_j \xrightarrow{D[0,1]} \sigma W(t), \quad \sigma^2 = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(Y_k,Y_0) < \infty; \tag{2.4}$$

(iii) the fourth-order cumulants satisfy the condition:

$$\sup_h \sum_{r,s=-\infty}^{\infty} |\mathrm{Cum}(Y_0,Y_h,Y_r,Y_s)| < \infty. \tag{2.5}$$

REMARK 2.1 Condition (iii) of Assumption A is required, in particular, for the consistency of the variance estimator (1.2). In fact, it is weaker than the usual condition in which $\sup_h$ is replaced by $\sum_h$, cf. Chapter 9 of Anderson (1971).

We now present two examples of processes satisfying Assumption A.

EXAMPLE 2.1 Linear processes $Y_k = \sum_{j=-\infty}^{\infty} a_j \varepsilon_{k-j}$, where the a_j are real weights, $\sum_j |a_j| < \infty$, and the ε_j are iid random variables with zero mean, unit variance and finite fourth moment $E\varepsilon_0^4 < \infty$ are well-known to satisfy Assumption A. The convergence of the finite dimensional distributions of $N^{-1/2}S_{[Nt]} = N^{-1/2}\sum_{j=1}^{[Nt]} Y_j$ can be shown, for example, similarly as in Ibragimov and Linnik (1971) (see also Giraitis and Surgailis (1986)), the tightness follows from the inequality $ES^4_{[Nt]} \le C[Nt]^2$, and condition (iii) from the formula

$$\mathrm{Cum}(Y_0,Y_h,Y_r,Y_s) = (E\varepsilon_0^4 - 3)\sum_{k=-\infty}^{\infty} a_k a_{k+h} a_{k+r} a_{k+s}.$$

EXAMPLE 2.2 Processes $\{\eta_k, k \in \mathbf{Z}\}$ satisfying the equations

$$\eta_k = \rho_k \xi_k, \quad \rho_k = a + \sum_{j=1}^{\infty} b_j \eta_{k-j}, \tag{2.6}$$

where the ξ_k are iid non-negative random variables with finite fourth moment $E\xi_0^4 < \infty$ and $a > 0$, $b_j \geq 0$, $j = 1, 2, \ldots$ arise in the study of the volatility of returns on speculative assets. For example, if $b_j = 0$ for $j > p$ and $\xi_k = \varepsilon_k^2$, $\varepsilon_k \sim N(0,1)$, then the η_j are the squares of random variables which follow the ARCH(p) model. The infinite sum in (2.6) allows to interpret the η_j as the squares, or any positive powers for that matter, of the more general GARCH(p, q) sequences, see Giraitis *et al.* (2000b) for more details and relevant references. Giraitis *et al.* (2000b) showed that if

$$(E\xi_0^4)^{1/4} \sum_{j=1}^{\infty} b_j < 1. \tag{2.7}$$

then the $Y_k = \eta_k - E\eta_k$ satisfy conditions (i), and validity of (iii) of Assumption A was established in Giraitis and Robinson (2001). Condition (ii) follows from Theorem 2.1 which we state here and whose proof is given in the Appendix.

THEOREM 2.1 *Suppose the η_k satisfy (2.6), $E\xi_0^4 < \infty$ and (2.7) holds. Then as $N \to \infty$*

$$N^{-1/2} \sum_{j=1}^{[Nt]} (\eta_j - E\eta_j) \overset{D[0,1]}{\longrightarrow} \sigma W(t),$$

where $\sigma^2 = \sum_{j=-\infty}^{\infty} \mathrm{Cov}(\eta_j, \eta_0)$.

We now state the assumptions A and B we impose on the trend $\{f^{(N)}(k)\}$ of the nonstationary series (2.1).

ASSUMPTION B. $\{f^{(N)}(k)\}_{k=1,\ldots,N}$, $N \geq 1$, is an array of real numbers for which there exists a positive sequence p_N and a function h on $[0, 1]$, which is not identically zero, such that

$$p_N^{-1} \sum_{k=1}^{[Nt]} f^{(N)}(k) \to h(t) \quad (N \to \infty) \tag{2.8}$$

and

$$\frac{p_N}{N^{1/2}} \to a^* \in [0, \infty].$$

The mode of convergence in (2.8) depends on the specific statistic. For the V/S (and KPSS) statistic, (2.8) has to hold in $L^2[0,1]$. For the R/S statistic, h is in $D[0,1]$ and (2.8) has to hold in the sup norm.

ASSUMPTION C. Assumption B is satisfied and there exists a positive sequence $r_N \to \infty$ and a number $0 < a < \infty$, such that as $N \to \infty$

$$r_N^{-1} \sum_{k=1}^{N} (f^{(N)}(k))^2 \to a, \tag{2.9}$$

$$\sum_{k=1}^{N-1} |f^{(N)}(k) - f^{(N)}(k+1)| k^{1/2} = O(r_N^{1/2}), \tag{2.10}$$

(2.11) $\sum_{k=1}^{v_N} |f^{(N)}(k)|^2 = o(r_N)$ for any $v_N = o(N)$, $\quad |f^{(N)}(k)|^2 = O(r_N/N)$ for $k \sim N$

and

$$\frac{p_N^2}{N r_N} \to b^* < \infty. \tag{2.12}$$

(Note that from (2.8), (2.9) it follows by the Cauchy inequality that $\limsup_N \dfrac{p_N^2}{N r_N} < \infty$.)

REMARK 2.2 Assumption B implies

$$p_N^{-1} \sum_{k=1}^{[Nt]} (f^{(N)}(k) - \bar{f}^{(N)}) \to h(t) - t\,h(1) =: h^0(t) \quad \text{in} \quad L^2[0,1]$$

and Assumption C implies

$$r_N^{-1} \sum_{k=1}^{N} (f^{(N)}(k) - \bar{f}^{(N)})^2 \to a - b^* h^2(1) =: \alpha, \tag{2.13}$$

where $\bar{f}^{(N)} = N^{-1} \sum_{k=1}^{N} f^{(N)}(k)$. Assumptions of this type with $p_N = r_N = N$ were used in Bhattacharya *et al.* (1983). Our assumptions looks somewhat more involved because we allow greater generality of nonstationarity/trend (for example, in the context considered by Bhattacharya *et al.* (1983) (2.12) holds trivially with $b^* = 1$).

We consider now some examples of trends.

EXAMPLE 2.3 (Change-point.) Suppose that

$$f^{(N)}(k) = \begin{cases} \mu_1 & \text{if} \quad 1 \le k \le k^*, \\ \mu_2 & \text{if} \quad k^* < k \le N, \end{cases}$$

where k^* denotes a change-point. We also assume that there is $0 < \tau^* < 1$ such that $k^* = [\tau^* N]$, and that $\Delta := \mu_1 - \mu_2 \neq 0$. Then Assumption B holds with $p_N = N$, $a^* = \infty$ and

$$h(t) = \mu_1 \; t \wedge \tau^* + \mu_2 (t - t \wedge \tau^*); \quad h^0(t) = \Delta \; (t \wedge \tau^* - t\tau^*).$$

Assumption C is satisfied with $r_N = N$, $b^* = 1$ and $a = \mu_1^2 \tau^* + \mu_2^2 (1 - \tau^*)$. Hence

$$\alpha = a - h^2(1) = \Delta^2 \tau^* (1 - \tau^*).$$

EXAMPLE 2.4 (Hyperbolic trend) Let

$$f^{(N)}(k) = c_1 |k + c_2|^\beta, \tag{2.14}$$

where $\beta > -1/2$; $c_1 \in \mathbf{R} \setminus \{0\}$, $c_2 \in \mathbf{R}$. This is the example considered by Bhattacharya *et al.* (1983). Assumption B is satisfied with $p_N = N^{1+\beta}$, $a^* = \infty$ and

$$h(t) = \frac{c_1 t^{\beta+1}}{\beta + 1}; \quad h^0(t) = h(t) - t h(1) = c_1 \; \frac{t(t^\beta - 1)}{\beta + 1}.$$

Assumption C is satisfied with $r_N = N^{1+2\beta}$, $b^* = 1$ and

$$a = \frac{c_1}{1 + 2\beta}; \quad \alpha = \frac{c_1^2}{1 + 2\beta} - \frac{c_1^2}{(1 + \beta)^2}.$$

Example 2.5 (Hyperbolic trend depending on N) Suppose that

$$f^{(N)}(k) = c_1|k + c_2N|^\beta, \tag{2.15}$$

where $\beta > -1/2$; $c_1 \in \mathbf{R} \setminus \{0\}$, $c_2 \in \mathbf{R}$. Then Assumption B is satisfied with $p_N = N^{1+\beta}$, $a^* = \infty$, and

$$h(t) = c_1 \int_0^t |x + c_2|^\beta dx; \quad h^0(t) = c_1 \left(\int_0^t |x + c_2|^\beta dx - t \int_0^1 |x + c_2|^\beta dx \right).$$

Assumption C is satisfied with $r_N = N^{1+2\beta}$, $b^* = 1$ and

$$\alpha = c_1^2 \left(\int_0^1 |x + c_2|^{2\beta} dx - \left(\int_0^1 |x + c_2|^\beta dx \right)^2 \right).$$

We now state the first main result of this section which allows us to analyse the impact of the nonstationarity (trend) in model $X_k = f^{(N)}(k) + Y_k$ with stationary short memory term $\{Y_k\}$ on R/S –type tests of long memory. We state it for the V/S and the modified R/S statistics $V_N/\hat{s}_{N,q}$ and $R_N/\hat{s}^2_{N,q}$ which are used in the long memory test below. The statements for other related statistics, like the KPSS statistic, are similar and can be readily obtained using Lemmas 6.2 and 6.3.

Theorem 2.2 *Suppose the X_k are given by (2.1) where the Y_k satisfy Assumption A and the $f^{(N)}(k)$ satisfy Assumption C. Consider the quantities*

$$Z_{a^*}(t) = \begin{cases} \sigma W^0(t) + a^* h^0(t) & \text{if} \quad a^* < \infty, \\ h^0(t) & \text{if} \quad a^* = \infty; \end{cases}$$

$$V(c^*) = \begin{cases} \alpha c^* + \sigma^2 & \text{if} \quad c^* < \infty, \\ \alpha & \text{if} \quad c^* = \infty; \end{cases}$$

$$T_N = \begin{cases} 1 & \text{if} \quad a^* < \infty, \\ p_N^2 N^{-1} & \text{if} \quad a^* = \infty; \end{cases} \qquad \widetilde{T}_N = \begin{cases} 1 & \text{if} \quad c^* < \infty, \\ \dfrac{N}{q_N r_N} & \text{if} \quad c^* = \infty. \end{cases}$$

Suppose $r_N \to \infty$, $q_N \to \infty$,$q_N/N \to 0$ and there exist the limit $q_N r_N N^{-1} \to c^* \in [0, \infty]$. Then

$$(T_N \widetilde{T}_N)^{-1} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \frac{\int_0^1 (Z_{a^*}(t))^2 dt - \left(\int_0^1 Z_{a^*}(t) dt \right)^2}{V(c^*)}$$

and

$$N^{1/2} (T_N \widetilde{T}_N)^{-1/2} \frac{R_N}{\hat{s}_{N,q}} \xrightarrow{d} \frac{\sup_{0 \le t \le 1} Z_{a^*}(t) - \inf_{0 \le t \le 1} Z_{a^*}(t)}{\sqrt{V(c^*)}}.$$

The proof of Theorem 2.2 is given in the Appendix.

Focusing on the V/S statistic, we now discuss the asymptotic behaviour of the test described in Definition 2.1 below in light of Theorem 2.2.

DEFINITION 2.1 (LM test): The null hypothesis of short memory is rejected in favour of a long memory alternative if $V_N/\hat{s}^2_{N,q} > K$, where $K > 0$ is a critical value.

Suppose first that there is no nonstationarity/trend ($f_k^{(N)} \equiv 0$). Giraitis *et al.* (2000a) showed that under the null hypothesis, more precisely for short memory process $\{Y_k\}$ satisfying Assumption A,

$$(2.16)\qquad \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \int_0^1 (W^0(t))^2 dt - \Big(\int_0^1 W^0(t)dt\Big)^2, \quad \text{as } q_N \to \infty,\ q_N/N \to 0,$$

where the limit distribution is parameter–free. In addition,under a long memory alternative which is precisely defined by Assumption L in Section 3, the test rejects the null hypothesis with power approaching 1.

Suppose now that the data contain a non-zero trend $f^{(N)}(k)$. Theorem 2.2 above shows that the presence of the trend in short memory data leads to the rejection of the short memory null hypothesis with power tending to 1 if

$$(2.17)\qquad T_N \widetilde{T}_N \to \infty.$$

For instance, (2.17) holds if $T_N = p_N^2/N$, $\widetilde{T}_N = \frac{N}{q_N r_N}$ and

$$(2.18)\qquad T_N\widetilde{T}_N \equiv \frac{p_N^2}{r_N q_N} \to \infty.$$

In particular, relation (2.18) holds true in Examples 2.3 and 2.4 – 2.5 (with $\beta \geq 0$), where $T_N\widetilde{T}_N = N/q_N \to \infty$. Hence, V/S and other R/S–type tests are sensitive both to the presence of long memory and a trend, in other words, they are prone to detect "spurious" long memory. Note that the trend in Theorem 2.2 is rather significant in the sense that $\|f^{(N)}\|_2 = (\sum_{k=1}^N (f^{(N)}(k))^2)^{1/2} \to \infty$ as $N \to \infty$ (see (2.9)). It is interesting to find out how small a trend can be so that it is no longer asymptotically detectable. With this question in mind, we shall now focus on *small* trends characterized roughly by the requirement that

$$(2.19)\qquad \limsup_N \|f^{(N)}\|_2 < C < \infty.$$

Condition (2.19) implies that Assumption B holds with $p_N = O(N^{1/2})$, and that (2.9) holds with $r_N = O(1)$ so that Assumption C cannot be satisfied with $r_N \to \infty$. Thus, the trends satisfying (2.19) fall outside the framework established in Lemma on p.653 of Bhattacharya *et al.* (1983). The following proposition shows that such small trends cannot, as a rule, be detected by the V/S test.

PROPOSITION 2.1 *Suppose that the X_k are given by* (2.1) *where the Y_k satisfy Assumption A and the $f^{(N)}(k)$ satisfy Assumption B with $p_N = o(N^{1/2})$,* (2.19) *holds and, in addition,*

$$(2.20)\qquad \sum_{k=1}^{N-1} |f^{(N)}(k) - f^{(N)}(k+1)| k^{1/2} = O(1),$$

$$(2.21)\qquad f^{(N)}(k) = O(N^{-1/2}) \quad \text{for} \quad k \sim N.$$

Then (2.16) *holds.*

PROOF: The proof of Lemma 6.3 below shows that under the assumptions of Proposition 2.1 $\hat{s}^2_{N,q} = \sigma^2 + O_P(q/N)$ which combined with Lemma 6.2 below implies (2.16). $\square$

EXAMPLE 2.6 *Small* trends given by (2.14), $f^{(N)}(k) = c_1|k+c_2|^\beta$, and by (2.15), $f^{(N)}(k) = c_1|k+c_2N|^\beta$, but with $\beta < -1/2$, have the property that $\limsup_N \|f^{(N)}\|_2 < \infty$ and satisfy the other assumptions of Proposition 2.1, so they cannot be detected by the LM test given in Definition 2.1 with power tending to 1.

3 Long memory and trend

To get a fuller picture, we look in this section at the asymptotic behaviour of the test statistics in situations when the long memory alternative is perturbed by a trend. We show that in this case $V_N/\hat{s}^2_{N,q} \xrightarrow{d} \infty$ no matter whether the trend is present or not, so that the short memory hypothesis is rejected with power tending to 1.

Denote by $W_H(t)$ a fractional Brownian motion with parameter H, i.e. a Gaussian process with mean zero and covariances $E[W_H(t_1)W_H(t_2)] = (1/2)(t_1^{2H} + t_2^{2H} - |t_1 - t_2|^{2H})$. Set also $W^0_{1/2+d}(t) = W_{1/2+d}(t) - tW_{1/2+d}(1)$. Giraitis *et al.* (2000a) showed that if there is no trend and the $X_k \equiv Y_k$ satisfy Assumption L below, then

$$\left(\frac{q_N}{N}\right)^{2d} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \int_0^1 (W^0_{1/2+d}(t))^2 dt - \Big(\int_0^1 W^0_{1/2+d}(t)dt\Big)^2. \tag{3.1}$$

Hence under the alternative of long memory $V_N/\hat{s}^2_{N,q} \xrightarrow{P} \infty$ at the rate $(N/q_N)^{2d}$, and so the short memory hypothesis will be rejected with power tending to 1.

In this section, we investigate how this rate and the asymptotic distribution are affected by the presence of a trend. We state the results only for the V/S statistic, as the conclusions for the R/S and KPSS statistics are the same and the corresponding formulas are easy to derive.

ASSUMPTION L. (Long memory) $\{Y_k\}$ is a fourth-order stationary sequence satisfying the following conditions:

(*i*)

$$\mathrm{Cov}(Y_k, Y_0) \sim ck^{2d-1}, \tag{3.2}$$

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

(*ii*)

$$\frac{1}{N^{1/2+d}} \sum_{j=1}^{[Nt]} (Y_j - EY_j) \xrightarrow{D[0,1]} c_d W_{1/2+d}(t), \tag{3.3}$$

where c_d is a positive number;

(*iii*)

$$\sup_h \sum_{r,s=-N}^{N} |\mathrm{Cum}(Y_h, Y_r, Y_s, Y_0)| = O(N^{2d}). \tag{3.4}$$

Obviously, under Assumption L, $\sum_{k=0}^\infty |\mathrm{Cov}(Y_k, Y_0)| = \infty$.

Before stating the results, we present two examples.

EXAMPLE 3.1 A standard class of processes satisfying Assumption L consists of linear sequences $Y_k = \sum_{j=0}^{\infty} a_j \varepsilon_{k-j}$, where the ε_k are iid random variables with zero mean, unit variance and finite fourth moment, and the coefficients a_j satisfy $a_j \sim cj^{d-1}, \quad j \to \infty$, for some $c > 0$ and $0 < d < 1/2$. It is easy to see that (3.2) holds. Convergence in (3.3) is well-known, see, e.g., Davydov (1970). Relation (3.4) is not difficult to verify, see Giraitis *et al.* (2000a).

EXAMPLE 3.2 A second important class of long memory processes, satisfying (3.2) and (3.3) consists of the long memory LARCH(∞) processes introduced by Robinson (1991) and investigated by Giraitis *et al.* (2000):

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

where $\{\varepsilon_k, \ k \in \mathbf{Z}\}$ is zero mean finite variance iid sequence, and α, β_j, $(\sum_{j=1}^{\infty} \beta_j^2 < \infty)$ are real coefficients. Assuming that $E\varepsilon_k^2 = 1$, σ_k^2 is seen to be the conditional variance of r_k. Then if β_j satisfy $\beta_j \sim cj^{d-1}$, $0 < d < 1/2$, the process $Y_k = r_k^2$ has long memory. In contrast to Example 2.2, neither α nor the β_j are assumed positive and σ_k, not its square, is a linear combination of the past r_k, rather than their squares. The r_k can be viewed as returns on a speculative asset which has "long memory in volatility". If $\sum_{j=1}^{\infty} \beta_j^2 < c$ for an appropriate constant c which depends on the fourth moment of the ε_k, then the $Y_k = r_k^2$ centered at the expectation satisfy conditions (i) and (ii) of Assumption L, see Giraitis *et al.* (2000). We conjecture that condition (iii) also holds, but this remains an open problem.

In the following theorem we consider the limiting distribution of the statistics $V_N/\hat{s}_{N,q}^2$ under trend and long memory.

THEOREM 3.1 *Suppose the X_k are given by (2.1) where the Y_k satisfy Assumption L and the $f^{(N)}(k)$ satisfy Assumption C. Let $r_N \to \infty$, $q_N \to \infty$, $q_N/N \to 0$ and there exists the limit $q_N^{1-2d} r_N N^{-1} \to c^* \in [0, \infty]$. Then*

$$(T_N \widetilde{T}_N)^{-1} \frac{V_N}{\hat{s}_{N,q}^2} \xrightarrow{d} \frac{\int_0^1 (Z_{a^*}(t))^2 dt - \left(\int_0^1 Z_{a^*}(t) dt \right)^2}{V(c^*)},$$

where

$$Z_{a^*}(t) = \begin{cases} c_d W_{1/2+d}^0(t) + a^* h^0(t) & \text{if} \quad a^* < \infty, \\ h^0(t) & \text{if} \quad a^* = \infty; \end{cases}$$

$$V(c^*) = \begin{cases} \alpha c^* + c_d^2 & \text{if} \quad c^* < \infty, \\ \alpha & \text{if} \quad c^* = \infty; \end{cases}$$

$$T_N = \begin{cases} N^{2d} & \text{if} \quad a^* < \infty, \\ p_N^2 N^{-1} & \text{if} \quad a^* = \infty; \end{cases} \qquad \widetilde{T}_N = \begin{cases} q_N^{-2d} & \text{if} \quad c^* < \infty, \\ \frac{N}{q_N r_N} & \text{if} \quad c^* = \infty; \end{cases}$$

(Notice that the quantities $T_N, \widetilde{T}_N, Z_{a^*}$ and $V(c^*)$ depend on d and are different from the corresponding quantities in Theorem 2.2.)

Theorem 3.1 shows that the short memory hypothesis for the model $X_k = f^{(N)}(k) + Y_k$ with a *large* trend $\{f_k^{(N)}\}$ and long memory process $\{Y_k\}$ will be rejected with power tending to 1 if $T_N \tilde{T}_N \to \infty$.

In the case of Examples 2.3 and 2.4–2.5 (with $\beta \geq 0$), it holds $T_N \tilde{T}_N = N/q_N$, so that by Theorem 3.1 the statistic $V_N/\hat{s}^2_{N,q}$ diverges to infinity at the rate N/q_N. This rate is the same as in Theorem 2.2 when $\{Y_k\}$ has short memory and it is faster than the rate $(N/q_N)^{2d}$ in (3.1) where (Y_k) has long memory and the trend is not present.

Theorem 3.1 follows from the following lemma whose proof is given in the Appendix.

LEMMA 3.1 *Under the assumptions of Theorem 3.1,*

$$T_N^{-1} V_N \xrightarrow{d} \int_0^1 (Z_{a^*}(t))^2 dt - \Big(\int_0^1 Z_{a^*}(t) dt \Big)^2, \tag{3.5}$$

and

$$\hat{s}^2_{N,q} = \frac{q_N r_N}{N}\ \alpha + q_N^{2d} c_d^2 + o_P\big(\frac{q_N r_N}{N} + q_N^{2d}\big). \tag{3.6}$$

The following proposition shows that *small* trends, such that $\limsup_N \|f^{(N)}\|_2 < \infty$, do not affect the limiting distribution of the V/S statistic (3.1). In this case the short memory hypothesis will be rejected and the rate of divergence of the statistics $V_N/\hat{s}^2_{N,q}$, $(N/q_N)^{2d}$, will be slower than N/q_N which occurs in the case of *large* trends and long memory.

PROPOSITION 3.1 *Suppose the X_k are given by* (2.1) *where the Y_k satisfy Assumption L and the $f^{(N)}(k)$ satisfy Assumption B with $p_N = o(N^{1/2})$ and* (2.19), (2.20) *and* (2.21). *If $q_N \to \infty$, $q_N/N \to 0$, then convergence* (3.1) *holds.*

PROOF: Use the same argument as in the proof of Proposition 2.1 and Theorem 3.1. □

4 The change–point model

In this section we study in greater detail Example 2.3. Our goal is to determine the order of the smallest break (jump) in the mean which can be registered by the test as a presence of long memory. We also discuss the possibility of developing a test procedure which, in case of a rejection of the null hypothesis of short memory, would allow to distinguish between the long memory and change-point alternatives. Again, we focus only on the V/S statistic.

Consider the general case where the magnitude of the jump depends on N:

$$f^{(N)}(k) = \begin{cases} \mu_1^{(N)} & \text{if} \quad 1 \leq k \leq k^*, \\ \mu_2^{(N)} & \text{if} \quad k^* < k \leq N, \end{cases} \tag{4.1}$$

where $k^* = [\tau^* N]$ and $0 < \tau^* < 1$. Set $\Delta_N := \mu_1^{(N)} - \mu_2^{(N)} \neq 0$, $h_*^0(t) := \tau^* \wedge t - \tau^* t$.

Proposition 4.1, whose proof is given in the Appendix, shows that the break in the mean of order Δ_N such that $\Delta_N / N^{-1/2} \to \infty$ leads to the rejection of the short memory hyphothesis (parts (i) and (iii)). On the other hand, a *small* break $\Delta_N = o(N^{-1/2})$ does not effect the limiting behaviour of the test (part (ii)).

Proposition 4.1 *Suppose the X_k and $f^{(N)}(k)$ are given by* (2.1) *and* (4.1), *respectively, the Y_k satisfy Assumption A and $\Delta_N N^{1/2} \to a^* \in [-\infty, \infty]$.*
(i) If $|\Delta_N| {q_N}^{1/2} \to \infty$ then

$$\frac{q_N}{N} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \frac{H^*}{\tau^*(1-\tau^*)} = \frac{\tau^*(1-\tau^*)}{12},$$

where

$$H^* := \int_0^1 (h^0_*(t))^2 dt - \Big(\int_0^1 h^0_*(t) dt\Big)^2 = \frac{(\tau^*(1-\tau^*))^2}{12};$$

(ii) if $\Delta_N = O(N^{-1/2})$ ($|a^*| < \infty$) then

$$\frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \sigma^{-2}\Big[\int_0^1 (\sigma W^0(t) + a^* h^0_*(t))^2 dt - \Big(\int_0^1 (\sigma W^0(t) + a^* h^0_*(t)) dt\Big)^2\Big].$$

(iii) if $|\Delta_N| N^{1/2} \to \infty$ ($|a^*| = \infty$) and $\Delta^2_N q_N \to c^*$ ($0 \le c^* < \infty$) then

$$\frac{q_N}{N} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \frac{c^* H^*}{c^* \tau^*(1-\tau^*) + \sigma^2} \qquad \text{if } 0 < c^* < \infty,$$

$$\frac{1}{(\Delta_N N^{1/2})^2} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \frac{H^*}{\sigma^2} \qquad \text{if } c^* = 0.$$

We now describe the asymptotic behaviour of the test for the observations $X_k = f^{(N)}(k) + Y_k$ with the change–point trend (4.1) and a long memory sequence Y_k. Similar arguments as in the proof of Lemma 3.1 and Proposition 4.1 yield

Proposition 4.2 *Suppose that the $f^{(N)}(k)$ are given by* (4.1) *and the Y_k satisfy Assumption L.*
(i) If $|\Delta_N| q_N^{1/2-d} \to \infty$ then

$$\frac{q_N}{N} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \frac{H^*}{\tau^*(1-\tau^*)} = \frac{\tau^*(1-\tau^*)}{12};$$

(ii) if $\Delta_N N^{1/2-d} \to a^* \in (-\infty, \infty)$ then

$$\Big(\frac{q_N}{N}\Big)^{2d} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} c_d^{-2}\Big[\int_0^1 (Z_{a^*}(t))^2 dt - \Big(\int_0^1 Z_{a^*}(t) dt\Big)^2\Big],$$

where $Z_{a^*}(t) = c_d W^0_{1/2+d}(t) + a^* h^0(t)$;
(iii) if $|\Delta_N| N^{1/2-d} \to \infty$, $\Delta^2_N q_N^{1-2d} \to c^* \in [0, \infty)$ then

$$\frac{q_N}{N} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \frac{c^* H^*}{c^* \tau^*(1-\tau^*) + c_d^2} \qquad \text{if } 0 < c^* < \infty,$$

$$\frac{1}{\Delta^2_N N^{1-2d}} \Big(\frac{q_N}{N}\Big)^{2d} \frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \frac{H^*}{c_d^2} \qquad \text{if } c^* = 0.$$

Statistic	Limit under assumption		
	Short memory	Long memory	Change point
$\dfrac{V_N}{\hat{s}^2_{N,q}}$	$U_{V/S}$	∞	∞
$\dfrac{q_N}{N}\dfrac{V_N}{\hat{s}^2_{N,q}}$	0	0	$\dfrac{\tau^*(1-\tau^*)}{12}$
$\left(\dfrac{q_N}{N}\right)^{2d}\dfrac{V_N}{\hat{s}^2_{N,q}}$	0	$Z_{V/S}$	∞

TABLE 1. The limit behaviour of the V/S statistics.

In the usual case magnitude of jump $\Delta_N = \Delta \neq 0$ in (4.1) does not depend on N, and Propositions 4.1 and 4.2 yield that

$$\frac{q_N}{N}\frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} \frac{\tau^*(1-\tau^*)}{12},$$

so the short memory hyphotheses will be rejected with the same rate N/q_N if the Y_k satisfy either the short memory Assumption A or the long memory Assumption L.

We conclude this section by summarizing in Table 1 the asymptotic behaviour of the statistic $V_N/\hat{s}^2_{N,q}$ under various normalizations. Recall that in the case where the $X_k \equiv Y_k$ are short-range dependent random variables satisfying Assumption A

$$V_N/\hat{s}^2_{N,q} \xrightarrow{d} U_{V/S} := \int_0^1 (W^0(t))^2 dt - (\int_0^1 W^0(t)dt)^2.$$

(If $F_{V/S}$ is the distribution of $U_{V/S}$ and F_K is the asymptotic distribution of the Kolmogorov statistic, then $F_{V/S}(x) = F_K(\pi\sqrt{x})$. Recall that the Kolmogorov statistic $\sqrt{N}\sup_{-\infty<t<\infty}|\hat{F}_N(t)-F(t)|$ determines the uniform distance between empirical and theoretical distribution functions. More details are presented in our paper Giraitis *et al.* (2000a).) In the case where the $X_k = Y_k$ satisfy Assumption L Theorem 3.1 implies that

$$\left(\frac{q_N}{N}\right)^{2d}\frac{V_N}{\hat{s}^2_{N,q}} \xrightarrow{d} Z_{V/S} := \int_0^1 (W^0_{1/2+d}(t))^2 dt - (\int_0^1 W^0_{1/2+d}(t)dt)^2.$$

As it was shown, the statistic $V_N/\hat{s}^2_{N,q}$ tends to infinity under both long memory (LM) and change–point (CP) alternatives, what means that both these models can explain the rejection of the null hypothesis of short memory. However, the statistic $(q_N/N)V_N/\hat{s}^2_{N,q}$ exhibits a different behaviour under assumptions LM and CP. Since under assumption CP it converges to a positive constant $\tau^*(1-\tau^*)/12$, this fact may allow (after construction of the corresponding second-order asymptotics) for large enough N to discriminate between a long memory process and a short memory process having a shift in mean (see the comment of Peter M. Robinson in Lobato and Savin (1998)). Obviously, to make the statistical difference between $\tau^*(1-\tau^*)/12$ and 0, the corresponding second–order asymptotics needs to be established. The practical recipes and simulation study are under development.

5 Conclusions

It has been realized for some time that tests for long memory are sensitive to the presence of nonstationarity (trend and change–point) in short memory data and as a rule "large" trends will be registered as long memory. Following the footsteps of of Bhattacharya *et al.* (1983), Künsch (1986) and Heyde and Dai (1996), among others, in this explanatory paper we developed a quantitative analysis of these phenomena in a broad setting which encompasses various types of trends and dependence structures.

The results of the present paper show that the rates of divergence to infinity of the test statistics (and the limits of the renormalized statistics) are different under short memory plus trend model and under the long memory, as well as under long memory plus trend and "pure" (no trend) long memory. The paper clarifies the difficulties of desriminating between long memory and several forms of nonstationarity, and it is hoped that a better understanding of the possible scenarios might aid the development of procedures aimed at discriminating between them. These procedures are likely to involve data driven algorithms for selecting q_N and would most likely rely on resumpling or subsampling techniques. The latter might be particularly relevant, because we established the existence of limiting distributions and rates of convergence, see Sections 3.5, 4.5 and 10.5.5 of Politis and Romano and Wolf (1999). Block bootstrap variance estimators, c.f. (1.4), might also be applicable, see e.g. Bühlmann and Künsch (1999) for a data driven procedure. Reseach of this type, which would require extensive numerical experimets, is however beyond the intended scope of the present paper.

6 Appendix

6.1 Proof of Theorem 2.1

As mentioned in Example 2.2, the $Y_j = \eta_j - E\eta_j$ satisfy conditions (2.3) and (2.5), cf., respectively, Proposition 3.1 in Giraitis *et al.* (2000b) and Corrolary 3.1 in Giraitis and Robinson (2001). These results combined with the convergence of the finite dimensional distributions allow us to use Lemma 6.1 below to provide a succinct proof of Theorem 2.1. We do not verify the convergence of the finite dimensional distributions here, as it can be established in a completely analogous way as the convergence of the one-dimensional distributions in Theorem 5.1 of Giraitis *et al.* (2000b). Lemma 6.1 may also be of independent interest.

LEMMA 6.1 *Suppose that a fourth-order stationary zero mean sequence* $\{Y_k\}$ *satisfies assumptions* (2.3), (2.5) *and*

$$N^{-1/2}\sum_{j=1}^{[Nt]} Y_j \xrightarrow{fdd} \sigma W(t). \tag{6.1}$$

Then weak convergence result (2.4) *holds.*

PROOF: Set

$$S_n := \sum_{k=1}^{n} Y_k, \quad n \geq 1.$$

Since, by (6.1), $N^{-1/2}S_{[Nt]} \stackrel{fdd}{\longrightarrow} \sigma W(t)$ it remains to check the tightness. For the proof of tightness it is sufficient to show that

$$ES_m^4 \leq Cm^2 \tag{6.2}$$

uniformly over $m = 0, ..., N$. Relation (6.2) implies that a tightness criterion for $N^{-1/2}S_{[Nt]}$ is satisfied (use a straightforward modification of Theorem 15.6 in Billingsley (1968)) Indeed, for any $t_1 < t < t_2$,

$$N^{-2}E\Big((S_{[Nt]} - S_{[Nt_1]})^2(S_{[Nt_2]} - S_{[Nt]})^2\Big) \leq N^{-2}\Big(E(S_{[Nt]} - S_{[Nt_1]})^4 E(S_{[Nt_2]} - S_{[Nt]})^4\Big)^{1/2}$$

$$\leq C\ N^{-2}([Nt] - [Nt_1])([Nt_2] - [Nt]) \leq CN^{-2}([Nt_2] - [Nt_1])^2.$$

It remains to prove (6.2). We have that

$$ES_m^4 = \sum_{j_1,\ldots,j_4=1}^{m} E(Y_{j_1}\cdots Y_{j_4}) = 3\sum_{j_1,\ldots,j_4=1}^{m} E(Y_{j_1}Y_{j_2})E(Y_{j_3}Y_{j_4}) + \sum_{j_1,\ldots,j_4=1}^{m} \mathrm{Cum}(Y_{j_1}, Y_{j_2}, Y_{j_3}, Y_{j_4})$$

using the relation

$$E(Y_{j_1}\cdots Y_{j_4}) = E(Y_{j_1}Y_{j_2})E(Y_{j_3}Y_{j_4}) + E(Y_{j_1}Y_{j_3})E(Y_{j_2}Y_{j_4})$$

$$+E(Y_{j_1}Y_{j_4})E(Y_{j_2}Y_{j_3}) + \mathrm{Cum}(Y_{j_1}, Y_{j_2}, Y_{j_3}, Y_{j_4})$$

which holds when $EY_k = 0$. Since

$$\mathrm{Cum}(Y_{j_1}, Y_{j_2}, Y_{j_3}, Y_{j_4}) = \mathrm{Cum}(Y_{j_1-j_4}, Y_{j_2-j_4}, Y_{j_3-j_4}, Y_0),$$

we get using assumptions (2.3), (2.5) that

$$ES_m^4 \leq 3(\sum_{j,k}^{m} E(Y_jY_k))^2 + m^2 \sup_h \sum_{j,k=-\infty}^{\infty} |\mathrm{Cum}(Y_j, Y_k, Y_h, Y_0)| \leq Cm^2.$$

□

6.2 Proof of Theorem 2.2 and Lemma 3.1

Theorem 2.2 follows immediately from Lemmas 6.2 and 6.3 below.

LEMMA 6.2 *Suppose that the X_k are given by* (2.1) *where the Y_k satisfy Assumption A and the $f^{(N)}(k)$ satisfy Assumption B. Then,*

$$T_N^{-1}V_N \stackrel{d}{\to} \int_0^1 (Z_{a^*}(t))^2 dt - \Big(\int_0^1 Z_{a^*}(t)dt\Big)^2 \tag{6.3}$$

and

$$N^{1/2}T_N^{-1/2}R_N \stackrel{d}{\to} \sup_{0\leq t\leq 1} Z_{a*}(t) - \inf_{0\leq t\leq 1} Z_{a*}(t). \tag{6.4}$$

Proof: We focus only on the verification of (6.3), for (6.4) an analogous argument is used.

Observe that

$$V_N = \sum_{k=1}^{N} S_{X,k}^2 - N^{-1}(\sum_{k=1}^{N} S_{X,k})^2$$

where

$$S_{X,k} = N^{-1}\sum_{j=1}^{k}(X_j - \bar{X}_N) = N^{-1}\sum_{j=1}^{k}(Y_j - \bar{Y}_N) + N^{-1}\sum_{j=1}^{k}(f^{(N)}(j) - \bar{f}_N)$$
$$=: S_{Y,k} + S_{f,k},$$

with $\bar{X}_N = N^{-1}\sum_{j=1}^{N} X_j$. Thus

$$V_N = \int_{1/N}^{1+1/N} \Big(N^{1/2} S_{Y,[Nt]} + (p_N N^{-1/2})(p_N^{-1} N) S_{f,[Nt]}\Big)^2 dt$$

$$-\Big(\int_{1/N}^{1+1/N} (N^{1/2} S_{Y,[Nt]} + (p_N N^{-1/2})(p_N^{-1} N) S_{f,[Nt]}) dt\Big)^2. \tag{6.5}$$

By Assumption A, as $N \to \infty$

$$N^{1/2} S_{Y,[Nt]} \stackrel{D[0,1]}{\longrightarrow} \sigma W^0(t)$$

and, by Assumption B,

$$p_N^{-1} N S_{f,[Nt]} \to h^0(t)$$

in $L^2[0,1]$ what together with (6.5) implies the statement of lemma. □

Lemma 6.3 *Suppose the X_k are given by (2.1) where the Y_k satisfy Assumption A and the $f^{(N)}(k)$ satisfy Assumption C. Let $r_N \to \infty$, $q_N \to \infty$, $q_N/N \to 0$. Then*

$$\hat{s}_{N,q}^2 = \frac{q_N r_N}{N}\,\alpha + \sigma^2 + o_P\left(\frac{q_N r_N}{N}\right) + o_P(1).$$

Proof: To estimate

$$\hat{s}_{N,q}^2 = \hat{\gamma}_0 + 2\sum_{j=1}^{q_N} \omega_j(q_N)\hat{\gamma}_j$$

write

$$\hat{\gamma}_j = N^{-1}\sum_{i=1}^{N-j}(X_i - \bar{X}_N)(X_{i+j} - \bar{X}_N) = \hat{\gamma}_j^{(1)} + \hat{\gamma}_j^{(2)} + \hat{\gamma}_j^{(3)},$$

where

$$\hat{\gamma}_j^{(1)} = N^{-1}\sum_{i=1}^{N-j}\nu_i\nu_{i+j}; \quad \hat{\gamma}_j^{(2)} = N^{-1}\sum_{i=1}^{N-j}(Y_i - \bar{Y}_N)(Y_{i+j} - \bar{Y}_N);$$

$$\hat{\gamma}_j^{(3)} = N^{-1}\sum_{i=1}^{N-j}[\nu_i(Y_{i+j} - \bar{Y}_N) + (Y_i - \bar{Y}_N)\nu_{i+j}]$$

and $\nu_j := f^{(N)}(j) - \bar{f}^{(N)}$. Thus

$$\hat{s}_{N,q}^2 = \hat{s}_{N,q;1}^2 + \hat{s}_{N,q;2}^2 + \hat{s}_{N,q;3}^2, \tag{6.6}$$

where

$$\hat{s}^2_{N,q;i} = \hat{\gamma}^{(i)}_0 + 2\sum_{j=1}^{q_N} \omega_j(q_N)\hat{\gamma}^{(i)}_j, \quad i = 1,2,3. \tag{6.7}$$

We show that, as $N \to \infty$,

$$N(r_N q_N)^{-1}\hat{s}^2_{N,q;1} \to \alpha, \tag{6.8}$$

$$\hat{s}^2_{N,q;2} \xrightarrow{P} \sigma^2 \tag{6.9}$$

and

$$N(r_N q_N)^{-1}\hat{s}^2_{N,q;3} \xrightarrow{P} 0, \tag{6.10}$$

what implies the statement of lemma.

We prove first (6.8). Choose $K > 0$ large. Since for j in (6.7) $Kj \le Kq_N = o(N)$, by assumption (2.11),

$$r_N^{-1}\sum_{i=1}^{jK} \nu_i^2 = o(1). \tag{6.11}$$

Write

$$\rho_N := |r_N^{-1}\sum_{i=1}^{N-j} \nu_i(\nu_{i+j} - \nu_i)| \le r_N^{-1}|\sum_{i=1}^{jK-1} \nu_i(\nu_{i+j} - \nu_i)|$$

$$+ r_N^{-1}|\sum_{i=jK}^{N-j} \nu_i(\nu_{i+j} - \nu_i)| =: \rho_{N,1} + \rho_{N,2}.$$

By (6.11),

$$\rho_{N,1} = o(1). \tag{6.12}$$

Consider the term $\rho_{N,2}$. Using summation by parts we get

$$\rho_{N,2} \le r_N^{-1}|\sum_{i=jK}^{N-j-1} (\nu_i - \nu_{i+1})\sum_{k=jK}^{i} (\nu_{k+j} - \nu_k)| + r_N^{-1}|\nu_{N-j}||\sum_{k=jK}^{N-j} (\nu_{k+j} - \nu_k)|.$$

We have that

$$|\sum_{k=jK}^{i} (\nu_{k+j} - \nu_k)| \le Cj^{1/2}r_N^{1/2}. \tag{6.13}$$

Indeed, by (2.13),

$$|\sum_{k=1}^{i}(\nu_{k+j} - \nu_k)| \le \sum_{k=1}^{j} |\nu_k| + \sum_{k=i+1}^{i+j} |\nu_k| \le 2j^{1/2}(\sum_{k=1}^{N} \nu_k^2)^{1/2} \le Cj^{1/2}r_N^{1/2},$$

which implies (6.13). By (2.11),

$$|\nu_{N-j}| \le C(r_N/N)^{1/2}$$

and therefore

$$\rho_{N,2} \le Cr_N^{-1/2}\sum_{i=jK}^{N-j-1} |\nu_i - \nu_{i+1}|j^{1/2} + Cr_N^{-1/2}|\nu_{N-j}|q_N^{1/2}$$

$$(6.14) \qquad \leq C r_N^{-1/2} K^{-1/2} \sum_{i=jK}^{N-j-1} |\nu_i - \nu_{i+1}| i^{1/2} + C(q_N/N)^{1/2} \leq CK^{-1/2} + C(q_N/N)^{1/2}.$$

The last inequality in (6.14) was obtained using (2.10). Since $q_N = o(N)$, $\rho_{N,2}$ can be made arbitrarily small choosing $K > 0$ large enough. This and (6.12) yield that $\rho_N = o(1)$.

Using (2.13), we obtain

$$N r_N^{-1} \hat{\gamma}_j^{(1)} = r_N^{-1} \sum_{i=1}^{N-j} \nu_i^2 + o(1) = r_N^{-1} \sum_{i=1}^{N} \nu_i^2 + o(1) = \alpha + o(1)$$

uniformly in $j = 1, \ldots, q_N$ because, by (2.11),

$$r_N^{-1} \sum_{i=N-j}^{N} \nu_i^2 \leq r_N^{-1} C(r_N/N) j \leq C(q_N/N) \to 0.$$

Hence

$$N(r_N q_N)^{-1} \hat{s}^2_{N,q;1} = q_N^{-1}(1 + 2\sum_{j=1}^{q_N} \omega_j(q_N))(\alpha + o(1)) \to \alpha$$

because

$$1 + 2\sum_{j=1}^{q_N} \omega_j(q_N) = 1 + q_N$$

and $q_N^{-1} \to 0$. This proves (6.8).

It was shown in Theorem 4.5 of Giraitis *et al.* (2000a) that under assumptions (2.3) and (2.5)

$$\hat{s}^2_{N,q;2} \xrightarrow{P} \sigma^2.$$

Therefore (6.9) holds.

We prove now (6.10). Using summation by parts it follows that

$$N|\hat{\gamma}_j^{(3)}| = \Big| \sum_{i=1}^{N-j} [\nu_i (Y_{i+j} - \bar{Y}_N) + \nu_{i+j}(Y_i - \bar{Y}_N)] \Big|$$

$$\leq \sum_{i=1}^{N-j-1} |\nu_i - \nu_{i+1}| | \sum_{k=1}^{i} (Y_{k+j} - \bar{Y}_N)| + |\nu_{N-j}| | \sum_{k=1}^{N-j} (Y_{k+j} - \bar{Y}_N)|$$

$$+ \sum_{i=1}^{N-j-1} |\nu_{i+j} - \nu_{i+j+1}| | \sum_{k=1}^{i} (Y_k - \bar{Y}_N)| + |\nu_N| | \sum_{k=1}^{N-j} (Y_k - \bar{Y}_N)|.$$

Under Assumption A (i),

$$E| \sum_{k=1}^{i} (Y_k - \bar{Y}_N)| \leq (E| \sum_{k=1}^{i} (Y_i - \bar{Y}_N)|^2)^{1/2} \leq C i^{1/2}; \quad E| \sum_{k=1}^{i} (Y_{k+j} - \bar{Y}_N)| \leq C i^{1/2}$$

uniformly in $i = 1, \ldots, N-j$ and $j = 1, \ldots, q_N$. Thus, by (2.10) and (2.11),

$$N r_N^{-1} E|\hat{\gamma}_j^{(3)}| \leq C r_N^{-1} \sum_{i=1}^{N-1} |\nu_i - \nu_{i+1}| i^{1/2} + C r_N^{-1}(|\nu_N| + |\nu_{N-j}|) N^{1/2} = O(r_N^{-1/2}) \to 0$$

uniformly in $j = 1, \dots, q_N$. Hence

$$N(r_N q_N)^{-1} E|\hat{s}^2_{N,q;3}| \le C q_N^{-1}(1 + 2\sum_{j=1}^{q_N} \omega_j(q_N)) o(1) \to 0.$$

This proves (6.10). □

Proof of Lemma 3.1: Relation (3.5) follows from Assumptions L and B, using the same argument as in the proof of Lemma 6.2. To show (3.6), use expression (6.6), $\hat{s}^2_{N,q} = \hat{s}^2_{N,q;1} + \hat{s}^2_{N,q;2} + \hat{s}^2_{N,q;3}$. By (6.8), $N(r_N q_N)^{-1}\hat{s}^2_{N,q;1} \xrightarrow{P} \alpha$. In Theorem 4.5 of Giraitis *et al.* (2000a) it was shown that under Assumption L, $q_N^{-2d}\hat{s}^2_{N,q;2} \to c_d^2$. The same argument as in the proof of Lemma 6.3 shows that

$$|\hat{s}^2_{N,q;3}| \le C\, \frac{N^d r_N^{1/2} q_N}{N} = C\left(\frac{r_N q_N}{N}\right)^{1/2}\left(\frac{q_N}{N}\right)^{1/2-d} q_N^d = \left(\frac{r_N q_N}{N}\right)^{1/2} o(1) q_N^d = o\left(\frac{r_N q_N}{N} + q_N^{2d}\right),$$

which implies (3.6). □

6.3 Proof of Proposition 4.1

It suffices to show that

$$T_N^{-1} V_N \xrightarrow{d} \int_0^1 Z_{a^*}(t)^2 dt - \Big(\int_0^1 Z_{a^*}(t) dt\Big)^2 \tag{6.15}$$

where

$$Z_{a^*}(t) = \begin{cases} \sigma W^0(t) + a^* h^0_*(t) & \text{if } |a^*| < \infty, \\ h^0_*(t) & \text{if } |a^*| = \infty; \end{cases} \qquad T_N = \begin{cases} 1 & \text{if } |a^*| < \infty, \\ \Delta_N^2 N & \text{if } |a^*| = \infty \end{cases}$$

and

$$\hat{s}^2_{N,q} = \sigma^2 + o_P(1), \quad \text{if } |a^*| < \infty; \tag{6.16}$$

$$\hat{s}^2_{N,q} = q_N \Delta_N^2 \tau^*(1-\tau^*) + \sigma^2 + o_P\left(q_N \Delta_N^2 \frac{q_N}{N} + [q_N \Delta_N^2 \frac{q_N}{N}]^{1/2} + 1\right), \quad \text{if } |a^*| = \infty. \tag{6.17}$$

Since $\bar{f}^{(N)} = \frac{k^*}{N}\Delta_N + \mu_2^{(N)}$ we get that the quantities $\nu_k = f^{(N)}(k) - \bar{f}^{(N)}$ are

$$\nu_k = \begin{cases} \Delta_N(1 - \frac{k^*}{N}) & \text{if } 1 \le k \le k^*, \\ -\Delta_N \frac{k^*}{N} & \text{if } k^* < k \le N. \end{cases} \tag{6.18}$$

Set $p_N = \Delta_N N$, $r_N = \Delta_N^2 N$.

From (6.18) it follows that

$$\sum_{k=1}^{[Nt]} \nu_k = \Delta_N\Big([Nt] \wedge k^* - [Nt]\,\frac{k^*}{N}\Big) = \Delta_N N\Big(t \wedge \tau^* - t\tau^* + O(N^{-1})\Big), \tag{6.19}$$

$$\sum_{k=1}^{N} \nu_k^2 = \Delta_N^2\left(\left(1 - \frac{k^*}{N}\right)^2 k^* + \left(\frac{k^*}{N}\right)^2 (N - k^*)\right) = \Delta_N^2 N\Big(\tau^*(1-\tau^*) + O(N^{-1})\Big). \tag{6.20}$$

Identity (6.19) and the proof of Lemma 6.2 imply (6.15).

Relations (6.16) and (6.17) follow by the same argument as in the proof of Lemma 6.3 using (6.20) and the following relations. Since $f^{(N)}(k) - f^{(N)}(k+1) = \nu_k - \nu_{k+1} = \Delta_N \mathbf{1}_{\{k=k^*\}}$, we have by (6.18)

$$\sum_{i=1}^{N-j} |\nu_i(\nu_{i+j} - \nu_i)| \le \Delta_N^2 q_N,$$

$$\sum_{k=1}^{N-1} |f^{(N)}(k) - f^{(N)}(k+1)|\sqrt{k} \le C|\Delta_N| N^{1/2}$$

and $\nu_k^2 = O(\Delta_N^2)$ as $k \sim N$. □

Acknowledgements. We would like to thank the referees for their useful comments and relevant references.

References

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

Baum, C. F., Barkoulas, J. T. and Caglayan, M. (1999). Long memory or structural breaks: can either explain nonstationary real exchange rates under the current float? *Journal of International Financial Markets, Institutions, and Money*, **9,** 359–376.

Bhattacharya, R. N., Gupta, V. K. and Waymire, E. (1983). The Hurst effect under trends. *Journal of Applied Probability*, **20,** 649–662.

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

Bos, C. S., Franses, P. H. and Ooms, M. (1999). Long memory and level shifts: re-analyzing inflation rates. *Empirical Economics*, **24,** 427–449.

Bühlmann, P. and Künsch, H. (1999). Block length selection in the bootstrap for time series. *Computational Statistics and Data Analysis*, **31,** 295–310.

Campbell, J. T., Lo, A. W, MacKinley, A. C. (1997). *The Econometrics of Financial Markets.* Princeton, New Jersey.

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

Diebold, F. X. and Inoue, A. (1999). Long memory and structural change. Preprint available at http://www.stern.nyu.edu/~fdiebold/papers/papers.html.

Giraitis, L., Kokoszka, P. and Leipus, R. (2000a). Rescaled variance and related tests for long memory in volatility and levels. Preprint available at http://www.math.usu.edu/~piotr/research.html.

Giraitis, L., Kokoszka, P. and Leipus, R. (2000b). Stationary ARCH models: dependence structure and Central Limit Theorem. *Econometric Theory*, **16,** 3–22.

Giraitis, L. and Robinson, P. (2001). Whittle estimation procedure of ARCH models. *Econometric Theory*, **17,** 608-631.

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

Giraitis, L. and Surgailis, D. (1986). Multivariate Appell polynomials and the central limit theorem. In: *Dependence in Probability and Statistics.* Birkhauser, Boston-Basel-Stuttgart.

Granger, C. W. J. and Hyung, N. (1999). Occasional structural breaks and long memory. Discusssion Paper 99-14. Department of Economics. University of California, San Diego.

Hall, P. (1997). Defining and measuring long-range dependence. *Fields Institute Communications*, **11,** 153–160.

Heyde, C. C. and Dai, W. (1996). On the robustness to small trends of estimation based on the smoothed periodogram. *Journal of Time Series Analysis*, **17,** 141–150.

Heyde, C. C. and Yang, Y. (1997). On defining long-range dependence. *Journal of Applied Probability*, **34,** 939–944.

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

Ibragimov, I. A. and Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters–Noordhoff, Groningen.

Künsch, H. (1986). Discrimination between monotonic trends and long-range dependence. *Journal of Applied Probability*, **23,** 1025–1030.

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.

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

Lobato, I. N. and 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. *Z. Wahrschein. verw. Gebiete*, **31,** 271–285.

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

Mandelbrot, B. B. and 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.

Mikosch, T. and Stărică, C. (1999). Change of structure in financial time series, long range dependence and the GARCH model. IWI-preprint 99–5–06.

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

Perron, P. and Vogelsang, T. (1992). Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business and Economic Statistics*, **10,** 301–320.

Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables. *Journal of Econometrics*, **80,** 355–385.

Politis, D. N., Romano, J. p. and Wolf, M. (1999). *Subsampling.* Springer-Verlag.

Robinson, P. M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics*, **47,** 67–84.

Teverovsky, V. and Taqqu, M. S. (1997). Testing for long–range dependence in the presence of shifting means or a slowly declining trend, using a variance–type estimator. *Journal of Time Series Analysis*, **18,** 279–304.

Willinger, W, Taqqu, M. S., Leland, W. E., Wilson D. E. (1995). Self-similarity in high-speed packet traffic: analysis and modeling of Ethernet traffic measurements. *Statistical Science*, **10,** 67–85.