COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale University
New Haven, Connecticut 06520

Cowles Foundation Discussion Paper No. 947

TESTING COVARIANCE STATIONARITY UNDER MOMENT CONDITION FAILURE WITH AN APPLICATION TO COMMON STOCK RETURNS

by

Peter C.B. Phillips

and

Mico Loretan

July 1990

TESTING COVARIANCE STATIONARITY UNDER MOMENT CONDITION FAILURE WITH AN APPLICATION TO COMMON STOCK RETURNS*

by

Peter C. B. Phillips

and

Mico Loretan

Cowles Foundation for Research in Economics
Yale University

Preliminary Draft: December 1989

Completed Version: June 1990

---

*We are most grateful to Adrian Pagan and Bill Schwert for supplying their stock return data used in Pagan and Schwert (1989a, 1989b). Our thanks also go to Glena Ames for wordprocessing and to the NSF for research support under Grant No. SES 8821180.

ABSTRACT

This paper studies tests for covariance stationarity under conditions which permit failure in the existence of fourth order moments. The problem is important because many econometric diagnostics such as tests for parameter constancy, constant variance and ARCH and GARCH effects routinely rely on fourth moment conditions. Moreover, such tests have recently been extensively employed with financial and commodity market data, where fourth moment conditions may well be quite tenuous and are usually untested. This paper considers several tests for covariance stationarity including sample split prediction tests, cusum of squares tests and modified scaled range tests. When fourth moment conditions fail we show how the asymptotic theory for these tests involves functionals of an asymmetric stable Levy process, in place of conventional standard normal or Brownian bridge asymptotics. An interesting outcome of the new asymptotics is that the power of these tests depends critically on the tail thickness in the data. Thus, for data with no finite second moment, the above mentioned tests are inconsistent. Some new tests for heterogeneity are suggested that are consistent in the infinite variance case. These are easily implemented and rely on standard normal asymptotics. A consistent estimator of the maximal moment exponent of a distribution is also proposed. Again this estimator is easily implemented, has standard normal asymptotics and leads to a simple test for the existence of moments up to a given order. An empirical application of these methods to the monthly stock return data recently studied in Pagan and Schwert (1989a, 1989b) and to daily returns of the Standard and Poors 500 stock index is presented.

JEL Classification Number: 211

Keywords: asymmetric stable process, characteristic exponent, covariance stationarity, cusum of squares test, maximal moment exponent, sample split prediction test, scaled range, stable Levy bridge, stock returns.

1. INTRODUCTION

An interesting feature of stock market data that has recently come to light is the apparent nonstationarity in the variance of stock returns. Pagan and Schwert (1989a, 1989b) present some strong evidence that stock return data over long periods cannot be assumed to be covariance stationary. In their (1989a) paper they plot recursive estimates of the variance of monthly stock returns from 1835–1987 and point to the dramatic increase in the variance in the decade after 1930. Figure 1 reproduces this plot of

(1) $$\hat{\mu}_2(t) = t^{-1}\Sigma_1^t(\hat{u}_k - \bar{u}_t)^2, \quad \bar{u}_t = t^{-1}\Sigma_1^t\hat{u}_k$$

where $\hat{u}_t$ is the difference between the actual stock return in period t and an estimate of its conditional mean (calculated by taking the residuals from a regression on monthly dummies, as described in §2.2 of Pagan and Schwert (1989a)). The distinctive pattern of this plot makes it highly unlikely that the data are covariance stationary. In their (1989b) paper, Pagan and Schwert formally test this hypothesis using a post–sample prediction test, a cusum test and a modified scaled range test. The results show strong evidence of departure from the null of covariance stationarity.

The idea of calculating the recursive variance (1) is due to Mandelbrot (1963), who suggested recursive variance plots as a useful diagnostic for the nonexistence of second order moments. When population variances do not exist, the behavior of the recursive sample variance is very different from that of a covariance stationary process. Figures 2 and 3 illustrate these differences by plotting recursive sample variances of sequences of iid draws from standard symmetric stable and Gaussian distributions, respectively. As is apparent from these plots the recursive sample variance of a stable process with exponent $\alpha < 2$ is subject to jumps and shows no tendency to settle down to a particular value as the sample size increases. In fact, for random draws from a stable law with exponent $\alpha < 2$ we have $\hat{\mu}_2(t) \to_{a.s.} \infty$. Such behavior is quite distinct from that of a Gaussian or other finite variance process where the recursive sample variance converges a.s. to the population variance as $t \to \infty$.

In the light of these differences it is natural to ask whether rejections of covariance stationarity for stock returns in formal statistical tests are simply a byproduct of thick tails in the generating mechanism. Most asymptotic distribution theory that is used in empirical econometric research relies on moment conditions that carefully control outlier occurrences. Indeed, it is not unusual, and in many cases quite reasonable, to see conditions of the type "let all required moments exist." However, in financial and commodity market time series the extent of the outlier activity casts doubt on the suitability of generic moment existence assumptions and this concern motivated Mandelbrot's original investigation. In the present case, it is important to note that since the recursive sample *variance* (1) is the object under study, conventional asymptotic distribution theory for this quantity calls for the existence of at least fourth order moments of the underlying data in the maintained hypothesis. This seems like a tall order when the series is common stock returns. Indeed, as we shall show below, the empirical evidence from the monthly stock return data used by Pagan and Schwert (1989a, 1989b) does not support this maintained hypothesis.

The present paper seeks to determine the effect on the asymptotic theory for statistics such as the recursive sample variance (1) when there is a relaxation in moment conditions on the underlying data. Our attention will concentrate on tests for covariance stationarity that involve sample second moments of the data. Principal among these are sample split prediction tests for constant variance, cusum of squares tests and modified scaled range tests. These tests were used by Pagan and Schwert (1989b) in their study of monthly stock returns and we will attempt to reevaluate their empirical findings in the light of the new asymptotic theory. There are many other applications of our theory that are relevant to diagnostic testing methods in econometrics, such as LM tests for ARCH and GARCH effects which also routinely rely on fourth moment conditions.

The paper is organized as follows. Section 2 presents some preliminary theory concerning the asymptotic distribution of sample second moments when fourth moments may not be finite. This theory is employed in Section 3 which develops an asymptotic theory for various tests of covariance stationarity. Conventional theory for these tests involves standard normal and Brownian bridge asymptotics. When fourth moment conditions fail we show that the new limit theory involves asymmetric stable processes and, in particular, a stable–Levy bridge process in place of the usual Brownian bridge and a quadratic variation process in place of the usual normalizing constant. It is also shown that conventional tests for homogenous variances have reduced asymptotic power when fourth moments are infinite and are actually inconsistent when the variance is infinite. We therefore propose a new test of heterogeneity that is consistent in the infinite variance case. This test involves the use of consistent estimates of the scale coefficient and characteristic exponent of the tail of a distribution of the asymptotic Pareto–Levy type. The estimate of the characteristic exponent may also be used to construct a consistent test about the size of the maximal moment exponent, for instance whether the fourth moment is finite. Section 4 describes the simulation methods employed in the computation of the asymmetric stable processes needed to tabulate and graph the functionals that represent the limit distributions in the case of infinite fourth moments. Comparisons with the conventional standard normal and Brownian bridge asymptotics are also displayed. Section 5 reports the empirical application to common stock returns. We use both the Pagan and Schwert data cited above, which is monthly data over the historical period 1834–1987, and daily stock return data over the recent cycle 1962–1987 from the Standard and Poors 500 series. Section 6 summarizes our main results and discusses some possibilities for future work.

2. PRELIMINARIES

Let (ϵ_t) be an iid sequence whose tail behavior is of the Pareto–Levy form, viz.

$$\text{(C1)} \qquad \begin{aligned} P(\epsilon > x) &= d_1 x^{-\alpha}(1 + \alpha_1(x)) , \; x > 0 , \; d_1 > 0 \\ P(\epsilon < -x) &= d_2 x^{-\alpha}(1 + \alpha_2(x)) , \; x > 0 , \; d_2 > 0 \end{aligned}$$

where $\alpha_i(x) \to 0$ $(i = 1, 2)$ as $x \to \infty$. When $0 < \alpha < 2$, (C1) ensures that ϵ lies in the normal domain of attraction of a stable law with characteristic exponent parameter α and we shall write $\epsilon \in \mathcal{ND}(\alpha)$ to signify this fact. When $\alpha > 2$, ϵ is in the normal domain of attraction of a normal distribution. In the latter case it is important to note that when $2 < \alpha < 4$ we have $\epsilon^2 \in \mathcal{ND}(\alpha/2)$, so that partial sums of ϵ^2 are no longer in the domain of attraction of a normal distribution. Obviously, such a distinction can play an important role in the asymptotic behavior of tests that are based on quantities like the recursive sample variance (1).

We add the following centering condition:

(C2) *If* $\alpha > 1$ *in* (C1) *then we require* $E(\epsilon) = 0$. *If* $\alpha = 1$ *we require* $\epsilon =_d -\epsilon$ (*i.e.* ϵ *is symmetrically distributed about the origin*).

Note that, when $\alpha < 1$ in (C1), no centering will be required. But this case is unlikely to be of importance in stock market data for which estimates of α have typically been in the range $1.2 < \alpha < 2.0$. (See Fama (1965) and more recent work by Blattberg and Gonedes (1974) and Fielitz and Rozelle (1982); see So (1987) on estimates of α for exchange rate series.) However, when $\alpha < 2$, no centering will be needed for partial sums of $\epsilon_t^2 \in \mathcal{ND}(\alpha/2)$ and this turns out to have very important implications on the asymptotic properties of tests based on (1), as we shall see in Section 3.

If the observed series is generated by the linear process

$$\text{(2)} \qquad y_t = \Sigma_{j=0}^{\infty} c_j \epsilon_{t-j}$$

where ϵ_t satisfies (C1) and (C2), then the series for y_t is convergent a.s. provided the coefficients c_j satisfy a suitable summability condition. We shall employ the following condition

(C3) $\Sigma_1^\infty j|c_j|^p < \infty$ for $0 < p < \alpha$, $p \leq 1$

because it is useful in the development of our asymptotic theory. Note that the series (2) converges a.s. provided $\Sigma_1^\infty |c_j|^p < \infty$ for $0 < p < \alpha$, $p \leq 1$ (c.f. Brockwell and Davis (1987), p. 480), so that (C3) is stronger than is necessary for (2) to be well defined. But note also that (C3) holds whenever y_t is generated by a stationary ARMA process because then the coefficients in (2) decline geometrically and thereby trivially satisfy (C3). Thus, (C3) is sufficiently general to cover most cases of interest.

If $\epsilon_t \in \mathcal{ND}(\alpha)$ with $0 < \alpha < 2$ we also have $y_t \in \mathcal{ND}(\alpha)$; and if ϵ_t follows a symmetric stable law then so too does y_t and we have the following distributional equivalence:

$$y_t =_d \left[\Sigma_{j=0}^\infty |c_j|^\alpha\right]^{1/\alpha} \epsilon_t .$$

For these and many other aspects of the theory of domains of attraction the reader is referred to the books by Brockwell and Davis (1987, Ch. 12.5) and by Ibragimov and Linnik (1971, Ch. 2). The limit theory for sample means and covariances of time series generated as in (2) was developed by Davis and Resnick (1985a, 1985b, 1986) and a recent treatment is given in Phillips and Solo (1989).

As remarked above, under (C3), (2) includes all stationary ARMA processes and in what follows it will be convenient for us to explicitly work with the AR(p) process

(3) $$y_t = \Sigma_{i=1}^p \varphi_i y_{t-i} + \epsilon_t$$

where the roots of $z^p - \Sigma_{i=1}^p \varphi_i z^{p-i} = 0$ all lie inside the unit circle.

Under (C1) and (C2) with $0 < \alpha < 2$ we have the normalizing sequence

(4) $$a_n = \inf\{x : P(|\epsilon| > x) \leq n^{-1}\} = dn^{1/\alpha}$$

for some constant d. When we parameterize the scale coefficients in (C1) as $d_1 = pa^{\alpha}$, $d_2 = qa^{\alpha}$ (with $p+q = 1$) we find that the constant coefficient in (4) is $d = a$ and then $a_n = an^{1/\alpha}$. With this normalization sequence (in place of the usual $n^{1/2}$ for finite variance models) we have the following limit laws for $0 < \alpha < 2$:

$$(5) \qquad a_n^{-1}\Sigma_1^n \epsilon_t \to_d U_\alpha(1)\,, \quad a_n^{-1}\Sigma_1^{[nr]} \epsilon_t \to_d U_\alpha(r)\,,$$

$$(6) \qquad \left[a_n^{-1}\Sigma_1^{[nr]}\epsilon_t,\, a_n^{-2}\Sigma_1^{[nr]}\epsilon_t^2\right] \to_d \left(U_\alpha(r),\, \textstyle\int_0^r (dU_\alpha)^2\right).$$

Here $U_\alpha(r)$ is the Levy α–stable process and $\int_0^r (dU_\alpha)^2 = [U_\alpha]_r$ is its quadratic variation process. The first result of (5) is classical (e.g. Ibragimov and Linnik (1971), Ch. 2), the second is its functional version, and (6) is a joint functional limit law for first and second sample moments that is proved in Resnick (1986, pp. 94–95). Some typical sample trajectories of a symmetric stable Levy process $U_\alpha(r)$ are plotted in Phillips (1990).

When $2 < \alpha < 4$ we have both $n^{-1}\Sigma_1^n \epsilon_t \to_{a.s.} 0$ and $n^{-1}\Sigma_1^n \epsilon_t^2 \to_{a.s.} \sigma_\epsilon^2 = E(\epsilon_t^2)$. However, since $\epsilon_t^2 \in \mathcal{ND}(\alpha/2)$ we also have a stable limit distribution theory for the sample second moments. In particular, we have

$$(7) \qquad a_n^{-2}\Sigma_1^n(\epsilon_t^2 - \sigma_\epsilon^2) \to_d U_{\alpha/2}(1)\,, \quad a_n^{-2}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_\epsilon^2) \to_d U_{\alpha/2}(r)$$

$$(8) \qquad \left[a_n^{-2}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_\epsilon^2),\, a_n^{-4}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_\epsilon^2)^2\right] \to_d \left[U_{\alpha/2}(r), \textstyle\int_0^r (dU_{\alpha/2})^2\right].$$

Since the distribution of $\epsilon_t^2 - \sigma_\epsilon^2$ is asymmetric with a finite left extremity $(-\sigma_\epsilon^2)$ the limit law represented by $U_{\alpha/2}(r)$ in (7) is an asymmetric stable process. The asymmetry coefficient in the limit stable law is given by $\beta = 1$. Moreover, as $\alpha \searrow 2$ the asymmetry of the limit process becomes more heavily accentuated. Ultimately, when $\alpha < 2$ we have

$$a_n^{-2}\Sigma_1^{[nr]}\epsilon_t^2 \to_d U^+_{\alpha/2}(r)$$

and in this case $U^+_{\alpha/2}(r)$ is a positive stable process on $D[0,1]$, i.e. the increments of $U^+_{\alpha/2}(r)$ are independent and follow a strictly positive stable law.

Finally, we observe that if y_t is generated by the AR(p) (3) the coefficients are consistently estimated by the OLS regression

(9) $$y_t = \Sigma_{i=1}^{p}\hat{\varphi}_i y_{t-i} + \hat{\epsilon}_t$$

irrespective of the value of α (e.g. see Kanter and Steiger (1974) and Hannan and Kanter (1977)). Correspondingly, the OLS residual $\hat{\epsilon}_t$ is consistent for ϵ_t for all $\alpha > 0$. When $\alpha > 2$, $\sigma_\epsilon^2 = E(\epsilon_t^2) < \infty$ and we have $\hat{\sigma}_\epsilon^2 = n^{-1}\Sigma_1^n\hat{\epsilon}_t^2 \to_p \sigma_\epsilon^2$. Moreover, the following limit theory applies to sample variances of the residuals $\hat{\epsilon}_t$.

LEMMA 2.1. *Let* (C1), (C2) *and* (3) *hold and* $\hat{\epsilon}_t$ *be the residuals from* (9).

(a) *If* $\alpha > 4$ *and* $v_\epsilon = E(\epsilon_t^2 - \sigma_\epsilon^2)^2 = \kappa_4 + 2\sigma^4$ *then*

$$(nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) \to_d B(r),$$

a standard Brownian bridge on $C[0,1]$.

(b) *If* $0 < \alpha < 4$ *then*

$$a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) \to_d U_{\alpha/2}(r) - rU_{\alpha/2}(1) = K_{\alpha/2}(r)$$

a stable–Levy bridge or tied down stable Levy process on $D[0,1]$. *When* $2 \le \alpha < 4$, $U_{\alpha/2}(r)$ *is a two tailed asymmetric stable process. When* $0 < \alpha < 2$, $U_{\alpha/2}(r)$ *is a positive stable process, i.e. a stable process whose increments are independent and follow a positive stable law with parameter* $\alpha/2$.

3. ASYMPTOTIC THEORY FOR TESTS OF COVARIANCE STATIONARITY

There are several ways of testing for homogeneous variances. We shall look first at the tests suggested in Pagan and Schwert (1989b) and later consider some others that look promising. Two general approaches are possible. The first is to work with the observed series y_t itself. For stock return data the series is nearly random but there is some evidence of moving average effects, which Pagan and Schwert (1989a) suggest may be due to nonsynchronous data and calendar effects. Thus, a second approach is to work with a general model like (2) and (3) that accommodates temporal dependence and use the residuals obtained from a suitable regression. Clearly, AR models offer a convenient choice for this approach and the OLS residuals $\hat{\epsilon}_t$ from regressions like (9) consistently estimate the time series innovations ϵ_t for all $\alpha > 0$.

We shall examine both these approaches below.

3.1. Sample Split Prediction Tests for Covariance Stationarity

Here we split the sample into two eras according to $n = n_1 + n_2$ with $n_1 = k_n n_2$ and consider the hypothesis that

$$H_0 : \ E\hat{\mu}_2^{(1)} = E\hat{\mu}_2^{(2)}$$

where

$$\hat{\mu}_2^{(1)} = n_1^{-1}\Sigma_1^{n_1} y_t^2 \,, \ \hat{\mu}_2^{(2)} = n_2^{-1}\Sigma_{n_1+1}^{n} y_t^2 \,.$$

Pagan and Schwert (1989b) set $k_n = 1$, define $\hat{\tau} = \hat{\mu}_2^{(1)} - \hat{\mu}_2^{(2)}$ and estimate the variance of $\hat{\tau}$ nonparametrically by $2\hat{v}$ where

$$\hat{v} = \hat{\gamma}_0 + 2\Sigma_{j=1}^{\ell}(1 - j/(\ell+1))\hat{\gamma}_j \,.$$

Here $\hat{\gamma}_j$ is the j'th serial covariance of y_t^2 and ℓ is a suitable lag truncation number ($\ell = 8$ in their empirical application). We assume that $\ell \to \infty$ as $n \to \infty$ in such a way that $\ell/n \to 0$. Then, when $\alpha > 4$, $\hat{v}$ is a consistent estimator of v by conventional theory under quite general conditions (cf. Andrews (1989), Newey and West (1987) and Phillips (1987)).

We shall consider the general case where $k_n \to k$ with $k > 0$. Then $k/(1+k)$ (respectively, $1/(1+k)$) is the fraction of the overall sample in the limit that is in the first (second) era. A t–ratio test statistic for a sample split in the variance between the two eras can then be constructed on the basis of anticipated normal asymptotics for $\hat{\tau}$. Indeed, when fourth moments are finite we have from the proof of part (a) of Theorem 3.1 below that

$$n_1^{1/2}\hat{\tau} \to_d N(0, (1+k)v) .$$

This leads to the statistic

$$V_k(\tau) = n_1^{1/2}\hat{\tau}/[(1 + k_n)\hat{v}]^{1/2} . \tag{10}$$

When $k = 1$, $V_1(\tau) = n_1^{1/2}\hat{\tau}/(2\hat{v})^{1/2}$ is the statistic suggested in Pagan and Schwert (1989b). When $k \neq 1$, $V_k(\tau)$ enables us to look at eras of unequal length.

The asymptotic distribution of $V_k(\tau)$ depends on the value of the parameter α as shown in the following:

THEOREM 3.1. *Assume* (C1)–(C3) *hold and* $k_n \to k$ *as* $n \to \infty$ *with* $k > 0$ *fixed. Then:*

(a) *If* $\alpha > 4$, $V_k(\tau) \to_d N(0,1)$.

(b) *If* $0 < \alpha < 4$,

$$V_k(\tau) \to_d \left[k\int_0^1 (dU_{\alpha/2})^2\right]^{-1/2} [(1+k)U_{\alpha/2}(k/(1+k)) - kU_{\alpha/2}(1)] = \bar{V}_k \text{ , say.} \tag{11}$$

In (b) $U_{\alpha/2}(r)$ *is an asymmetric stable process with characteristic exponent* $\alpha/2$ *when* $2 \leq \alpha < 4$ *and a positive stable process when* $0 < \alpha < 2$.

REMARKS

(i) When $\alpha > 4$, the limit distribution of $V_k(\tau)$ is standard normal, so that conventional critical values from the $N(0,1)$ distribution can be used in testing H_0 under this maintained hypothesis that moments of the data are finite up to at least the fourth order.

(ii) When $\alpha < 4$ the limit distribution of $V_k(\tau)$ is a ratio of correlated stable variates. The limit distribution simplifies considerably when the length of the eras is the same in the limit, i.e. $k_n \to 1$.

COROLLARY 3.2. *Suppose* $k = 1$ *and* $0 < \alpha < 4$ *. Then*

$$V_1(\tau) \to_d \left[\int_0^1 (dU^s_{\alpha/2})^2\right]^{-1/2} U^s_{\alpha/2}(1), \tag{12}$$

where $U^s_{\alpha/2}(r)$ *is a symmetric stable process with characteristic exponent* $\alpha/2$ *on* $D[0,1]$.

The limit distribution given by (12) is, in fact, the same as the limit distribution of a self normalized sum (or t–ratio) formed from an i.i.d. sequence of variates in the domain of attraction of a stable law with exponent parameter $\alpha/2$. Logan *et al.* (1973), Resnick (1986), and Phillips (1990) give various representations of this limit distribution—see Phillips (1990) equation (46) for the representation given here. It is known to be bimodal (Logan *et al.* (1973) and Phillips and Hajivassiliou (1987) provide graphical plots), and nominal critical values from the $N(0,1)$ distribution are known to be conservative at the conventional levels 1% and 5%, but to lead to reductions in power (cf. Efron (1968)). Thus, when $\alpha < 4$ we can expect that tests based on $V_1(\tau)$ and standard normal critical values will suffer power reductions compared with the case where $\alpha > 4$. This will be explored further below when we consider the limit behavior of power functions of this test.

(iii) When $\alpha > 4$, $U_{\alpha/2}(r)$ is standard Brownian motion $W(r)$ and since $(dW)^2 = dr$ a.s. we have

$$\int_0^1 (dU_{\alpha/2})^2 = 1 \text{ a.s.}$$

for the denominator of (11). A small calculation shows that

$$k^{-1/2}[W(k/(1+k)) - k\{W(1) - W(k/(1+k))\}] =_d N(0,1)$$

so that the limit result (11) also yields part (a) as $\alpha \nearrow 4$.

(iv) When $k \neq 1$ and $\alpha < 4$, the limit distribution of $V_k(\tau)$ depends on k . This is because the construction of $V_k(\tau)$ is based on the explicit behavior of $n^{1/2}\hat{\tau}$ when $\alpha > 4$ – see (10) – and ignores the random limit of the denominator of $V_k(\tau)$ when $\alpha < 4$. The use of subsample estimates also leads to a dependence on k in the limit. To see this, let $\hat{v}_1$ and $\hat{v}_2$ be estimates of v constructed just as $\hat{v}$ but based on the two separate data sets $\{y_t^2 : t = 1, \ldots, n_1\}$ and $\{y_t^2 : t = n_1 + 1, \ldots, n\}$. The composite estimate of v is now

$$\tilde{v} = \hat{v}_1 + k_n \hat{v}_2 .$$

When $\alpha > 4$ we have $\hat{v}_1, \hat{v}_2 \to_p v$ and, hence, $\tilde{v} \to (1+k)v$, the same limit as that of $\hat{v}$. Use of the composite variance estimate $\tilde{v}$ leads to the modified t–ratio statistic

$$\tilde{V}_k(\tau) = n_1^{1/2} \hat{\tau} / \tilde{v}^{1/2} .$$

The asymptotic distribution of $\tilde{V}_k(\tau)$ is as follows:

THEOREM 3.3. *Under the same conditions and with the same notation as Theorem* 3.1 *we have:*

(a) *if* $\alpha > 4$, $\tilde{V}_k(\tau) \to_d N(0,1)$;

(b) *if* $0 < \alpha < 4$

$$\tilde{V}_k(\tau) \to_d \left[\int_0^{k/(1+k)} (dU_{\alpha/2})^2 + k^2 \int_{k/(1+k)}^1 (dU_{\alpha/2})^2\right]^{-1/2} \times [U_{\alpha/2}(k/(1+k)) - k\{U_{\alpha/2}(1) - U_{\alpha/2}(k/(1+k)\}] \tag{13}$$

$$\tilde{V}_1(\tau) \to_d \left[\int_0^1 (dU^s_{\alpha/2})^2\right]^{-1/2} U^s_{\alpha/2}(1) . \tag{14}$$

In (14) $U^s_{\alpha/2}(r)$ *is a symmetric stable process with characteristic exponent* $\alpha/2$ *as in Corollary* 3.2.

Analogous results to those of Theorem 3.1 and 3.2 apply to split sample tests that are based on the residuals $\hat{\epsilon}_t$ from the fitted model (9). We shall work below with the analogue of $\tilde{V}_k(\tau)$. Define

$$\hat{\tau}_\epsilon = n_1^{-1}\Sigma_1^{n_1}\hat{\epsilon}_t^2 - n_2^{-1}\Sigma_{n_1+1}^n\hat{\epsilon}_t^2 = \hat{\sigma}_{1\epsilon}^2 - \hat{\sigma}_{2\epsilon}^2 ,$$

$$\hat{v}_{1\epsilon} = n_1^{-1}\Sigma_1^{n_1}(\hat{\epsilon}_t^2 - \hat{\sigma}_{1\epsilon}^2)^2 , \ \hat{v}_{2\epsilon} = n_2^{-1}\Sigma_{n_1+1}^n\left[\hat{\epsilon}_t^2 - \hat{\sigma}_{2\epsilon}^2\right]^2 , \ \hat{v}_\epsilon = n^{-1}\Sigma_1^n\left[\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2\right]^2 ,$$

$$\tilde{v}_\epsilon = \hat{v}_{1\epsilon} + k_n\hat{v}_{2\epsilon} ,$$

and construct the t–statistics

$$V_k(\tau_\epsilon) = n_1^{1/2}\hat{\tau}_\epsilon/\hat{v}_\epsilon^{1/2} , \ \tilde{V}_k(\tau_\epsilon) = n_1^{1/2}\hat{\tau}_\epsilon/\tilde{v}_\epsilon^{1/2} .$$

The asymptotic theory for these statistics is as follows:

THEOREM 3.4. *If* (3) *is the generating mechanism for* y_t *and if* (C1)–(C2) *hold, then* $V_k(\tau_\epsilon)$ *(respectively,* $\tilde{V}_k(\tau_\epsilon)$ *) is asymptotically equivalent to* $V_k(\tau)$ $(\tilde{V}_k(\tau))$ *and has the same limit theory as that given in Theorem* 3.1 (3.3).

To examine the consistency of these tests we relax H_0. Let $(\epsilon_t)_{-\infty}^{\infty}$ be split into the two half series of iid variates $(\epsilon_t)_{-\infty}^{n_1}$ and $(\epsilon_t)_{n_1+1}^{\infty}$ that individually follow (C1) but with different scale coefficients. For the finite variance case $(\alpha > 2)$ we shall employ

$$H_1 : \sigma_\epsilon^2 \neq \sigma_{\epsilon+}^2$$

where $\sigma_\epsilon^2 = E(\epsilon_t^2)$ for $t \leq n_1$ and $\sigma_{\epsilon+}^2 = E(\epsilon_t^2)$ for $t > n_1$. Under H_1 we have

$$E(y_t^2) = \sigma_y^2 = (\Sigma_0^\infty c_j^2)\sigma_\epsilon^2 , \ t \leq n_1$$

$$E(y_t^2) = \sigma_{ys}^2 = (\Sigma_{j=0}^{s-1}c_j^2)\sigma_{\epsilon+}^2 + (\Sigma_{j=s}^\infty c_j^2)\sigma_\epsilon^2 , \ t = n_1 + s > n_1 .$$

For the infinite variance case $(\alpha < 2)$ we introduce heterogeneity through the scale coefficient in the tail of the distributions, as prescribed by (C1). We do so by using the scale coefficients (d_1, d_2) for $(\epsilon_t)_{-\infty}^{n_1}$ and (d_1^+, d_2^+) for $(\epsilon_t)_{n_1+1}^\infty$. If we assume that ϵ_t is symmetrically distributed then we can write $d_1 = d_2 = (1/2)a^\alpha$, $d_1^+ = d_2^+ = (1/2)a_+^\alpha$. The alternative hypothesis is then simply expressed as

$H_1' : a \neq a_+$.

The following result gives the properties of the covariance stationarity tests $V_k(\tau)$, $\tilde{V}_k(\tau)$ and $\tilde{V}_k(\tau_\epsilon)$ under these alternatives, H_1 and H_1' .

THEOREM 3.5. *Assume* (C1)–(C3) *hold and* $k > 0$. *Two cases apply.*

(a) $\alpha > 2$: *Under* H_1 , *the statistics* $V_k(\tau)$ *and* $\tilde{V}_k(\tau)$ *diverge as* $n \to \infty$ *and tests based on these statistics are consistent. Specifically, the rates of divergence as* $n \to \infty$ *are given by:*

(15) $$V_k(\tau), \tilde{V}_k(\tau) = O_p(n^{1/2}) , \text{ for } \alpha > 4$$

(16) $$V_k(\tau), \tilde{V}_k(\tau) = O_p(n^{1-2/\alpha}) , \text{ for } 2 < \alpha \leq 4 .$$

(b) $0 < \alpha < 2$: *Under* H_1' , *tests based on the statistics* $V_k(\tau)$ *and* $\tilde{V}_k(\tau)$ *are inconsistent. Specifically, as* $n \to \infty$:

(17) $$V_k(\tau), \tilde{V}_k(\tau) = O_p(1) .$$

If (C1)–(C2) *and* (3) *hold then identical results apply for the tests based on the statistic* $V_k(\tau_\epsilon)$ *and* $\tilde{V}_k(\tau_\epsilon)$.

REMARKS

(i) The $V_k(\tau)$ and $V_k(\tau_\epsilon)$ tests are consistent provided $\alpha > 2$. But note from (16) that the rate of divergence of $V_k(\tau)$ under the alternative slows as $\alpha \searrow 2$. We can therefore expect the power properties of the test to be very unsatisfactory when α is close to 2.

(ii) When $0 < \alpha < 2$ the tests are inconsistent. Thus, all of these tests have no real discriminatory power in identifying heterogeneity in the sample observations under conditions of infinite variance. If we seek to determine whether heterogeneity is present in such cases, other tests are required.

3.2. The Cusum of Squares Test for Covariance Stationarity

This test is based on the cumulative sums of $y_t^2 - \hat{\mu}_2$ where $\hat{\mu}_2 = n^{-1}\Sigma_1^n y_t^2$, leading to the statistic

$$\psi_n(r) = (n\hat{v})^{-1/2}\Sigma_1^{[nr]}(y_t^2 - \hat{\mu}_2) .$$

Alternatively, we can use deviations of the squared residuals, $\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2$, leading to the statistic

$$\psi_n^\epsilon(r) = (n\hat{v}_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) .$$

Both $\psi_n(r)$ and $\psi_n^\epsilon(r)$ are studentized cusum of squares statistics. In this sense they differ from the original cusum of squares statistic suggested in Brown, Durbin and Evans (1975). The original cusum of squares statistic is known to be not robust to departures from normality and its asymptotic distribution is sensitive to fourth moments. This is usually overcome by estimating fourth moments and studentizing the statistic, as suggested in Ploberger and Kramer (1986). The resulting statistic is entirely analogous to $\psi_n^\epsilon(r)$. The statistic $\psi_n(r)$ is similar in form but involves the estimate $\hat{v}$ of the "long–run" fourth moment of the data (i.e. the spectrum of y_t^2 at the origin) rather than simply the fourth moment of y_t itself. Pagan and Schwert (1989b) employ $\psi_n(r)$ in their empirical work.

Sample realizations of $\psi_n(r)$ and $\psi_n^\epsilon(r)$ lie in the function space $D[0,1]$ and a limit distribution theory must be worked out by using suitable weak convergence methods on this space. For models where y_t^2 has finite variance $(\alpha > 4)$ this presents no difficulty and the limit process of $\psi_n(r)$ and $\psi_n^\epsilon(r)$ is a standard Brownian bridge, whose sample paths lie almost surely in $C[0,1]$. When $\alpha < 4$ the limit process is different, is no longer confined to $C[0,1]$ and weak convergence in $D[0,1]$ does not always obtain in the J_1–Skorohod topology due to possible serial dependence in y_t^2 (see Avram and Taqqu (1986, 1989a)). For this reason, it is especially convenient to work with the cusum statistic $\psi_n^\epsilon(r)$ that is based on the squared residuals $\hat{\epsilon}_t^2$ from the autoregression (9). For, under

(3), $\hat{\epsilon}_t$ is consistent to ϵ_t and is thereby serially independent asymptotically. In this case, weak convergence of $\psi_n^{\epsilon}(r)$ does apply in D[0,1] and $\psi_n^{\epsilon}(r)$ can form the basis of a suitable cusum test.

We give the following limit theory.

THEOREM 3.6. *Assume* (C1)–(C3) *hold. Then:*

(a) *If* $\alpha > 4$

(18) $$\psi_n(r),\ \psi_n^{\epsilon}(r) \to_d B(r)$$

a standard Brownian bridge on C[0,1] .

(b) *If* $0 < \alpha < 4$

(19) $$\psi_n(r) \to_{fdd} L_{\alpha/2}(r) = K_{\alpha/2}(r) / \left[\int_0^1 (dU_{\alpha/2})^2\right]^{1/2}$$

(20) $$\psi_n^{\epsilon}(r) \to_d L_{\alpha/2}(r) = K_{\alpha/2}(r) / \left[\int_0^1 (dU_{\alpha/2})^2\right]^{1/2}$$

where $K_{\alpha/2}(r) = U_{\alpha/2}(r) - rU_{\alpha/2}(1)$ *is a stable–Levy bridge on* D[0,1] .

REMARKS

(i) When $\alpha > 4$, the limit distribution of both cusums $\psi_n(r)$ and $\psi_n^{\epsilon}(r)$ is the standard Brownian bridge. Pagan and Schwert (1989b) use bands that are based on critical values of the finite dimensional distributions, viz. $B(r) =_d N(0, r(1-r))$ for inference in graphical plots of their cusum statistic $\psi_n(r)$. These bands differ from those originally envisaged by Brown *et al.* (1975) for the cusum of squares statistic and by Durbin (1969) for the accumulated periodogram. In these papers the bands are designed so that the probability that the statistic hits the barrier at some point in its trajectory is controlled at the size of the test. This means that the probability that the trajectory is ever on or beyond the barrier corresponds to the level of the test. The situation is quite different for the finite dimensional distributions (fdd) bands that are based on critical values of $B(r) =_d N(0, r(1-r))$ for fixed r . In this case, the probability that a sample trajectory lies outside the fdd bands is greater than the nominal size, leading to a liberal test. This is because, for example, $P[B(r) > c] \leq P[\sup_s B(s) > c]$ for all r. The fdd bands do tell us something: if the sample trajectory lies inside the bands then non rejection is certainly the right decision. But these fdd bands do understate the extremes of sample trajectories and thereby lead to size distortions in testing stationarity by overrejection under the null.

(ii) When $\alpha < 4$ the limit theory is quite different. First, for $\psi_n(r)$, only the finite dimensional distributions converge (written as " $\to_{fdd}$ " in (19)) when the underlying data is serially dependent and we cannot in general assert that the random function $\psi_n(r)$ converges in $D[0,1]$. As shown in Avram and Taqqu (1989b), serial dependence in the process leads to successive jumps in the trajectories of the process which usually prevent convergence of partial sum processes like $\psi_n(r)$ in the J_1–Skorohod topology. This means that mass of the distribution escapes as the partial sums fluctuate too wildly for the sequence of probability measures associated with $\psi_n(r)$ to be tight. For this reason, it seems inappropriate to use $\psi_n(r)$ as a statistic for testing covariance stationarity. However, the statistic $\psi_n^{\epsilon}(r)$ that is based on regression residuals does converge in $D[0,1]$ when $0 < \alpha < 4$. The limit process (20) is a ratio of the stable–Levy bridge on $D[0,1]$ represented by $K_{\alpha/2}(r)$ and the correlated, positive stable process represented by the multiple stochastic integral $\int_0^1 (dU_{\alpha/2})^2$. Bands that are based on critical values of the finite dimensional distributions of $L_{\alpha/2}(r)$ and its extrema $\sup_r$, $\inf_r L_{\alpha/2}(r)$ can be used for inference in the cusum plot for $\psi_n^{\epsilon}(r)$. They are computed and applied in Sections 4 and 5 below.

(iii) The stable process $U_{\alpha/2}(r)$ that appears in (20) is asymmetric for $2 \leq \alpha \leq 4$ and strictly positive when $0 < \alpha < 2$. Examples of the sample trajectories of such processes and the associated bridge process $K_{\alpha/2}(r)$ are illustrated in Figures 4a and 4b. The distribution of the bridge process $K_{\alpha/2}(r)$ is skew symmetric in the sense that

$$K_{\alpha/2}(r) =_d - K_{\alpha/2}(1-r) .$$

To prove this we write

$$\begin{aligned}
K_{\alpha/2}(r) &= U_{\alpha/2}(r) - rU_{\alpha/2}(1) = \int_0^r dU_{\alpha/2} - r\int_0^1 dU_{\alpha/2} \\
&= \int_0^1 dU_{\alpha/2} - \int_r^1 dU_{\alpha/2} - r\int_0^1 dU_{\alpha/2} \\
&= (1-r)\int_0^1 dU_{\alpha/2} - \int_r^1 dU_{\alpha/2} \\
&= -\left\{\int_r^1 dU_{\alpha/2} - (1-r)\int_0^1 dU_{\alpha/2}\right\} \\
&=_d -\left\{\int_0^{1-r} dU_{\alpha/2} - (1-r)\int_0^1 dU_{\alpha/2}\right\} \\
&= -\left\{U_{\alpha/2}(1-r) - (1-r)U_{\alpha}(1)\right\} \\
&= -K_{\alpha/2}(1-r)\ .
\end{aligned}$$

Note that the skew symmetry of $K_{\alpha/2}(r)$ implies that of the limit process $L_{\alpha/2}(r)$ in (20) and this is reflected in the confidence contours for this process that we compute in Section 4 below. Note also that the skew symmetry of $K_{\alpha/2}(r)$ also implies the following distributional equivalence

$$\bar{V}_k =_d -\bar{V}_{1/k}$$

for the limit variate of the $V_k(\tau)$ statistic given in Theorem 3.1. This latter property generalizes the result given earlier in Corollary 3.2 that $\bar{V}_k$ is symmetric when $k = 1$.

(iv) Consistency properties of the cusum of square tests can be studied in the same way as the sample split prediction tests. Since $\psi_n^{\epsilon}(r)$ has a functional limit law for all $\alpha > 0$ we focus on the power properties of this test below. It is seen that the rates of divergence are comparable with those of the sample split prediction tests given in Theorem 3.5. In particular, the cusum test has decreasing power as $\alpha \searrow 2$ and is inconsistent for $0 < \alpha < 2$.

THEOREM 3.7. *Suppose* (C1)–(C2) *and* (3) *hold. Then we have:*

(a) $\alpha > 2$: *Under* H_1 *tests based on* $\psi_n^\epsilon(r)$ *are consistent with the following rates of divergence:*

(21) $$\psi_n^\epsilon(r) = O_p(n^{1/2}) \,, \text{ for } \alpha > 4$$

(22) $$\psi_n^\epsilon(r) = O_p(n^{1-2/\alpha}) \,, \text{ for } 2 < \alpha \leq 4 \,.$$

(b) $0 < \alpha < 2$: *Under* H_1' , *tests based on* $\psi_n^\epsilon(r)$ *are inconsistent and we have* $\psi_n^\epsilon(r) = O_p(1)$.

3.3. The Modified Scaled Range Test

This test is based on the extent of the observed maximum fluctuation in the cusum of squares statistic. We define

$$R_n = \sup_r \psi_n(r) - \inf_r \psi_n(r)$$

$$R_n^\epsilon = \sup_r \psi_n^\epsilon(r) - \inf_r \psi_n^\epsilon(r) \,.$$

These are functionals of $\psi_n(r)$ and $\psi_n^\epsilon(r)$ on $D(0,1]$. Using Theorem 3.6 and the continuous mapping theorem (both sup and inf are continuous functionals in the J_1–Skorohod topology) we get

(a) for $\alpha > 4$

$$R_n, R_n^\epsilon \rightarrow_d \sup_r B(r) - \inf_r B(r) = R_B \,, \text{ say}$$

(b) for $0 < \alpha < 4$

$$R_n^\epsilon \rightarrow_d \sup_r L_{\alpha/2}(r) - \inf_r L_{\alpha/2}(r) = R_L \,, \text{ say.}$$

Critical values for R_B are tabulated e.g. in Haubrich and Lo (1988, table 1a) and critical values for R_L are tabulated in Table 4 below. Observe that there is no corresponding limit for R_n in case (b) since the random element $\psi_n(r)$ is not weakly convergent in $D[0,1]$ in the J_1–Skorohod topology, as earlier discussed.

Using Theorem 3.4 we deduce that tests based on R_n^{ϵ} are consistent when $\alpha > 2$ but inconsistent when $\alpha < 2$. Tests based on R_n are consistent when $\alpha > 4$ but inconsistent when $\alpha < 4$, due to the failure of R_n to converge in this case.

3.4. A Consistent Test for Heterogeneity in the Infinite Variance Case

When $0 < \alpha < 2$ all of the statistics considered earlier lead to inconsistent tests under heterogeneity of the data. Since the data has infinite variance when $0 < \alpha < 2$ it might be argued that there is little point in testing for heterogeneity using sample variances. This is certainly borne out by Theorems 3.5 and 3.7. However, if the focus of interest is the more general hypothesis of stationarity (rather than constant variance or covariance stationarity) and it is suspected that this hypothesis breaks down over the sampling period, then we would expect such a breakdown to become evident in some characteristics of the data, if not the sample variances.

One way to test H_1' under $0 < \alpha < 2$ is to work directly with consistent estimates of the scale coefficients that appear in (C1). These may be constructed using order statistics in the following manner. Let $(\hat{\epsilon}_t)_1^n$ be the residuals from (9) and let $\hat{\epsilon}_{n1} \leq \hat{\epsilon}_{n2} \leq \cdots \leq \hat{\epsilon}_{nn}$ be the order statistics corresponding to this sample of residuals. Next, define

$$\hat{\alpha}_s = \left[s^{-1}\Sigma_{j=1}^{s} \ln \hat{\epsilon}_{n,n-j+1} - \ln \hat{\epsilon}_{n,n-s}\right]^{-1} \tag{23}$$

$$\hat{d}_s = sn^{-1}(\hat{\epsilon}_{n,n-s})^{\hat{\alpha}_s} \tag{24}$$

for some integer s. It is assumed that n is large enough and s/n small enough so that $\hat{\epsilon}_{n,n-s} > 0$ and thus $\hat{\alpha}_s$ and $\hat{d}_s$ are well defined real quantities. These estimators were originally suggested by Hill (1975) as conditional maximum likelihood estimators of the characteristic exponent parameter α and the scale coefficient $d = d_1$ in (C1). The asymptotic theory for them in the general case of a distribution whose tails have the asymptotic Pareto–Levy form (C1) is due to Hall (1982), who shows that it is optimal, at least in terms of the asymptotic bias and variance of these estimates, to choose the integer $s = s(n)$ so that it tends to infinity with n and is of order $n^{2\gamma/(2\gamma+\alpha)}$ when $\alpha_i(x) = O(x^{-\gamma})$ in C(1).

There is some advantage to choosing $s(n)$ so that $s/n^{2\gamma/(2\gamma+\alpha)} \to 0$ as $n \to \infty$. For, in this case we have from Theorem 2 of Hall (1982)

$$(25) \qquad s^{1/2}(\hat{\alpha}_s - \alpha) \to_d N(0, \alpha^2)$$

and

$$(26) \qquad s^{1/2}[\ell n\,(n/s)]^{-1}(\hat{d}_s - d) \to_d N(0, d^2)\,.$$

These asymptotics apply at a slightly reduced rate, viz. $s^{1/2}$ and $s^{1/2}[\ell n\,(n/s)]^{-1}$ rather than the rates $n^{\gamma/(2\gamma+\alpha)}$ and $n^{\gamma/(2\gamma+\alpha)}[\ell n\,(n)]^{-1}$ which apply when $s(n) = O(n^{2\gamma/(2\gamma+\alpha)})$. But they have the advantage that the limit distributions (25) and (26) involve only scale nuisance parameters which are easily eliminated in statistical tests.

Suppose, for example, we wish to mount a sample split prediction test using (26). We split the sample into two eras, $(\hat{\epsilon}_t)_1^{n_1}$ and $(\hat{\epsilon}_t)_{n_1+1}^{n}$ of approximately equal length and let $\hat{d}_s^{(1)}$, $\hat{d}_s^{(2)}$ be the corresponding scale coefficient estimators obtained by applying (24) to each of these subsamples. Next define the deviation

$$\tau_d = \hat{d}_s^{(1)} - \hat{d}_s^{(2)}$$

and the t–statistic

$$(27) \qquad \tilde{V}(\tau_d) = s^{1/2}[\ell n\,(n_1/s)]^{-1}\tau_d\Big/\left[\hat{d}_s^{(1)^2} + \hat{d}_s^{(2)^2}\right]^{1/2}.$$

The asymptotic theory for $\tilde{V}(\tau_d)$ yields a standard $N(0,1)$ tests. We have:

THEOREM 3.8. *Let* (C1), (C2) *and* (3) *hold and suppose* $k_n \to k = 1$ *as* $n \to \infty$. *Assume that* $\alpha_i(x) = O(x^{-\gamma})$ *for* $i = 1, 2$ *in* (C1) *for some* $\gamma > 0$. *Let* $s \to \infty$ *and* $s/n_1^{2\gamma/(2\gamma+\alpha)} \to 0$ *as* $n_1 \to \infty$. *Then, for all* $\alpha > 0$,

(28) $$\tilde{V}(\tau_d) \to_d N(0,1) .$$

REMARKS

(i) Tests for heterogeneity in the sample that are based on $\tilde{V}(\tau_d)$ may be applied easily using critical values from the $N(0,1)$ distribution. In constructing the statistic (27), however, a choice must be made in selecting s, the number of order statistics that are employed in formulae (23) and (24). If the tails are strictly Pareto (with $\gamma = \infty$) then $s = o(n_1)$ will suffice in Theorem 3.8. If the tails are not strictly Pareto (which seems more likely) but are well represented by an asymptotic series in $x^{-\alpha}$, then $\gamma = \alpha$ and we require $s = o(n_1^{2/3})$. In a situation of ignorance about the form of the tails of the distribution it seems appropriate to obtain estimates of the tail scale parameter d for a range of possible values of s. These can be centered on the value $s = n_1^{2/3}/\ell n(\ell n\, n_1)$, which would be appropriate in the case of a tail with an asymptotic series representation in power of $x^{-\alpha}$. The sensitivity of the statistic $\tilde{V}(\tau_d)$ and thereby the outcome of the test to the selection of s can then be examined directly in a given data set.

(ii) Note that the standard $N(0,1)$ asymptotics (28) apply for all values of $\alpha > 0$. Thus, the test may be applied irrespective of the value of the tail exponent α in (C1) and therefore provide an alternative to the sample split prediction tests $V(\tau)$ and $V(\tau_\epsilon)$. However, when $\alpha > 4$ the order statistic test $V(\tau_\alpha)$ will have lower power than $V(\tau)$ and $V(\tau_\epsilon)$. Indeed, rates of divergence under H_1' are given in the following result.

THEOREM 3.9. *Let the same conditions as those of Theorem* 3.8 *apply. Then, under* H_1', *the test based on* $\tilde{V}(\tau_d)$ *is consistent for all* $\alpha > 0$. *Specifically, as* $n \to \infty$ *we have the following rate of divergence under* H_1' :

(29) $\tilde{V}(\tau_d) = O_p(s^{1/2}/\ln(n_1/s))$.

(iii) Suppose we have $s = n_1^{2/3}/\ln(\ln n_1)$, as suited for tails with an asymptotic series representation in power of $x^{-\alpha}$. This choice leads to a divergence rate under H_1' of $O_p(n^{1/3}/\ln(n)(\ln \ln(n))^{1/2})$. Comparing these results with (16) we see that the order statistic test $\tilde{V}(\tau_d)$ will have greater asymptotic power whenever $\alpha < 3$. Thus, the test statistic $\tilde{V}(\tau_d)$ would seem to be worth using even in the finite variance case provided the tails are not too thin.

(iv) Observe that tests based on $\tilde{V}(\tau_d)$ may be applied to both the right and left tails of the distribution. In the latter case we simply use the order statistics for the alternate series $\eta_t = -\hat{\epsilon}_t$.

(v) Subsample tests of heterogeneity may also be constructed for the characteristic exponent α . In this case we may employ estimates of α based on the use of (23) for the two eras. Let these estimates be $\hat{\alpha}_s^{(1)}$ and $\hat{\alpha}_s^{(2)}$, respectively. Again we shall assume that the eras are of approximately equal length. Define

$$\tau_\alpha = \hat{\alpha}_s^{(1)} - \hat{\alpha}_s^{(2)}$$

and construct the t–statistic

$$\tilde{V}(\tau_\alpha) = s^{1/2}\tau_\alpha/\left[\hat{\alpha}_s^{(1)^2} + \hat{\alpha}_s^{(2)^2}\right]^{1/2} .$$

This leads to a test of the hypothesis

$$H_0^\alpha : \alpha^{(1)} = \alpha^{(2)} = \alpha$$

that the characteristic exponents $\alpha^{(1)}$ and $\alpha^{(2)}$ in (C1) are the same in the two eras. Again, the test is an asymptotic $N(0,1)$ test and is easy to implement. We have:

THEOREM 3.10. *Let the same conditions as those of Theorem* 3.8 *apply. Then under the null hypothesis* H_0^α *we have*

(30) $\tilde{V}(\tau_\alpha) \to_d N(0,1)$

and under the alternative hypothesis $H_1^{\alpha} : \alpha^{(1)} \neq \alpha^{(2)}$ *we have*

$$\tilde{V}(\tau_{\alpha}) = O_p(s^{1/2}) \tag{31}$$

and tests based on $\tilde{V}(\tau_{\alpha})$ *are consistent.*

3.5. Testing Hypotheses about the Maximal Moment Exponent

The maximal moment exponent of a distribution whose tails satisfy (C1) is given by the parameter α since $\alpha = \sup_q \{E|\epsilon|^q < \infty\}$. Of course, α is unknown in applications and it will often be useful to combine estimation of α with a test of the hypothesis that it has a particular value, say

$$H_0 : \alpha = \alpha_0$$

against the one sided alternative that

$$H_1 : \alpha < \alpha_0 .$$

Of prominent interest will be the two cases $\alpha_0 = 4$ and $\alpha_0 = 2$. For then, the alternative H_1' corresponds to the cases of moment condition failure that we have studied earlier where different asymptotic theory comes into play. A suitable test is based on the studentized statistic

$$V(\alpha) = s^{1/2}(\hat{\alpha}_s - \alpha_0)/\hat{\alpha}_s ,$$

or its LM version

$$V_0(\alpha) = s^{1/2}(\hat{\alpha}_s - \alpha_0)/\alpha_0 .$$

Under H_0 we have, from (25), the same limit theory for both statistics, viz.

$$V(\alpha), V_0(\alpha) \to_d N(0,1) ,$$

leading to one tailed tests based on the standard normal distribution. These tests are easy to apply and may form part of the preliminary diagnostic checking of the data characteristics. We shall illustrate their use in Section 5.

4. SIMULATIONS, CRITICAL VALUES AND GRAPHICAL ILLUSTRATIONS

In the previous section, we obtained representations for the asymptotic distributions of the sample split prediction test, the studentized cusum of squares test, and the modified scaled range test for covariance stationarity. For $\alpha < 4$, the limit laws of these statistics depend on functionals of stable processes. Closed form expressions for the probability density functions of these laws are unknown and we resort to Monte Carlo simulation to characterize their properties and to obtain appropriate critical values. Of particular interest is the extent to which the new distributions differ from those that apply in the standard case of finite fourth moments $(\alpha > 4)$. Only cases of $\alpha > 2$ will be considered here, since for $\alpha < 2$ the tests are inconsistent and thus of no interest to empirical research. To perform the simulations, we generate stable random variates and from these construct sample trajectories of the appropriate stable processes. Exact algorithms for generating stable random numbers have been proposed by Kanter and Steiger (1974) for the symmetric case and by Chambers, Mallows and Stuck (1976)[1] for the general asymmetric case. We considered values of $\alpha = \{2.1, 2.5, 3.0, 3.5, 3.8\}$, and set $n = 1{,}000$ as our "large" sample size, except for the asymmetric case when $\alpha = 2.1$, where we set $n = 2{,}500$. We performed 50,000 iterations of all experiments. To increase efficiency, the symmetry and skew–symmetry of the distributions was exploited in computing the critical values and densities presented here. The simulations were carried out in the GAUSS programming language; copies of these programs are available from the authors on request.

We first study the large sample distribution of the sample split prediction test statistic $V_1(\tau)$. We performed the following simulation experiment: We drew n iid

[1]We used their formula (2.3). We did not use the algorithm proposed in (4.1), which is based on a modified skewness parameter β', since we are only interested in generating maximally skewed stable variates here, and (2.3) is faster to compute than (4.1). Note that when $1 < \alpha < 2$ one must set the skewness parameter $\beta = -1$ in (2.3) to obtain stable variates with maximal *positive* skewness.

symmetric stable random numbers x_i of index $\alpha/2$, set $y_i = n^{-2/\alpha}x_i$ (note that by self-similarity $y_i \approx_d dU^s_{\alpha/2}$) and computed the ratio $(\Sigma_i^n y_i)/(\Sigma_i^n y_i^2)^{1/2}$, thus simulating the distribution of $U^s_{\alpha/2}(1)/(\int_0^1 (dU^s_{\alpha/2})^2)^{1/2}$. Critical values, at the usual levels of significance, are shown in Table 1 , together with the standard normal critical values which are applicable when $\alpha > 4$. As remarked in Section 3.1, the new critical values for typical test sizes are all lower than the conventional ones. For a two-sided test with size 5% , say, the applicable critical value declines from 1.96 (for $\alpha > 4$) to 1.73 (for $\alpha = 2.1$). Thus, use of the conventional critical values in cases where the true α is less than 4 leads to conservative tests. An estimate of the density of the $V_1(\tau)$ statistic for the case of $\alpha = 3$, computed from the simulations using a normal kernel, is graphed in Figure 5 . The density is quite different from the standard normal density: it is platykurtic, its tails are thinner than those of the normal distribution, and the density is bimodal, with peaks at -1 and $+1$ and a fairly flat region at the origin. Further Monte Carlo-based estimates of densities of t-ratio statistics when $\alpha/2 < 2$ are given in Phillips and Hajivassiliou (1987). Logan *et al.* (1973) computed the asymptotic densities of the t-ratio statistic when $\alpha/2 < 2$ through numerical integration of the associated characteristic functions.

We turn to the empirical distribution of the limit law $L_{\alpha/2}(r)$ of the cusum of squares statistics $\psi_n(r)$ and $\psi_n^\epsilon(r)$. We drew n iid *asymmetric* stable variates x_i of index $\alpha/2$, set $y_i = n^{-2/\alpha}x_i$, and computed $(\Sigma_i^{[nr]} y_i - r \cdot \Sigma_i^n y_i)/(\Sigma_i^n y_i^2)^{1/2}$ as the large sample approximation to $(U_{\alpha/2}(r) - r \cdot U_{\alpha/2}(1))/(\int_0^1 (dU_{\alpha/2})^2)^{1/2}$, for $r = \{0.1, 0.2, \ldots, 0.9\}$. The resulting critical values for typical test sizes as well as the median of the finite dimensional distributions (fdd) of $L_{\alpha/2}(r)$ are given in Tables 2a-2e; the exact critical values for the fdd of the Brownian Bridge process $B(r)$ are shown in Table 2f for comparison. Only the upper confidence levels are provided, the lower confidence contours being obtained from the skew-symmetric relationship $L_{\alpha/2}(r) =_d -L_{\alpha/2}(1-r)$. In Figures 6a-6f we graph the upper and lower confidence contours corresponding to (two–sided) 95% and 99% confidence levels. These tables and figures present a complex picture, which we shall discuss in steps. First, whereas the fdd confidence contours of the process B(r) are symmetric (Table 2.f and Figure 6.f), the contours become increasingly asymmetric as $\alpha \downarrow 2$. For $\alpha < 4$, the upper contours increase rapidly with r for small r, and return to 0 only slowly as $r \uparrow 1$. Next, the medians of the fdd's of $L_{\alpha/2}(r)$ also depend on r: the median is negative for $r < 0.5$ and positive for $r > 0.5$. The fdd's of $L_{\alpha/2}(r)$ as a function of α, for $\alpha = \{2.1, 3.0, 4.0\}$, are further contrasted in Figures 7.a and 7.b, in which we graph the 97.5% and 99.5% (one–sided) upper confidence contours, respectively. Third, for $\alpha < 4$, tests based on the (nominal $\alpha > 4$) *upper* fdd critical values are conservative for $r \geq 0.5$, but become increasingly liberal as $r \downarrow 0$. For $r = 0.5$, e.g., the 2–sided 99 % critical value decreases from 1.29 (the $\alpha > 4$ case) to 1.19 $(\alpha = 3)$ and further to 0.96 $(\alpha = 2.1)$, so that the conventional critical value leads to a conservative test in the latter cases. But for $r = 0.1$, the 99% upper critical value *increases* from 0.77 $(\alpha > 4)$ to 0.99 $(\alpha = 3)$ and 0.98 $(\alpha = 2.1)$, whereas the corresponding lower critical value *decreases* (in absolute value) from –0.77 $(\alpha > 4)$ to –0.58 $(\alpha = 3)$ and –0.36 $(\alpha = 2.1)$. Figure 8 summarizes the dependence of the shape of the upper confidence contours of the fdd's on α in a three–dimensional graph, for $2.25 \leq \alpha \leq 4$.

To complement the information on the fdd critical values given in Tables 2a–2e, we give the asymptotic critical values for the statistic $\sup_r(\psi_n(r))$, delivered from simulating $\sup_r(L_{\alpha/2}(r))$, in Table 3. Note that $P(\inf_r(L_{\alpha/2}(r) < c)$ $= P(\sup_r(L_{\alpha/2}(r) > -c)$. As discussed in Section 3.3 above, use of these critical values does not lead to the size distortions that result from looking only at the critical values of the fdd of $L_{\alpha/2}(r)$. The dependence of the critical values of $\sup_r(L_{\alpha/2}(r))$ on α is easily described: they decrease monotonically with α, so that use of $\alpha > 4$ based critical values will again lead to conservative tests. Evaluation of an empirical cusum of squares statistic is best based on both criteria, with the cusum showing behavior throughout the sample as well as points of maximum deviation, and critical values for the latter being delivered by the $\sup_r(L_{\alpha/2}(r))$ and $\inf_r(L_{\alpha/2}(r))$ statistics. Further, α is usually not known in empirical work and must be estimated in advance; if α cannot be estimated with high precision, the cusum of squares statistic should be evaluated using the critical values for a range of values of α around the point estimate.

Finally, critical values of the scaled range statistic $R_L = \sup_r(\psi_n(r)) - \inf_r(\psi_n(r))$ are given in Table 4, together with critical values of the conventional statistic R_B. As for the statistic $\sup_r(L_{\alpha/2}(r))$ above, critical values of R_L decrease with α, so that conventional critical values are conservative when the true values of α is less than 4. For a two–sided test of size 5%, say, the appropriate critical value decreases from 1.86 $(\alpha > 4)$ to 1.71 $(\alpha = 3)$ to 1.41 $(\alpha = 2.1)$.

5. ARE STOCK MARKET RETURNS COVARIANCE STATIONARY?

We start by making some general observations about the relevance of covariance stationarity to theory and empirical research. Constancy of the unconditional second moments of stock returns is rarely implied by models of optimizing behavior of economic agents, and tests for covariance stationarity may therefore appear to be of little importance to economic and financial theory. Indeed, the efficient market hypothesis is typically formulated in terms of restrictions that are placed on the first moment or expected value of the data, e.g., due to the intervention of arbitrage traders, "excess" returns cannot be earned, at least not on the basis of public information alone. In spite of the apparent focus of theory models on first moment behavior we still expect to have some prior information about second moments. For instance, we may reasonably expect the unconditional variance of stock market returns *not* to be constant over long periods of time. The weights on the individual stocks that enter the market index change over time; if the individual stocks have differing intrinsic volatilities, the changing weights will obviously affect the variance of the market portfolio. Also, institutional innovations in the stock markets, such as the rising importance of mutual funds and stock index arbitrage, seem to affect the variance of returns significantly and irreversibly. Notwithstanding these reasonable empirical expectations, the assumption of covariance stationarity is useful and convenient in applied and theoretical time series analysis. Indeed, it seems that covariance stationarity is often assumed more for statistical convenience than for theoretical economic reasons. Further, formal models of conditional heteroskedasticity such as those in Engle (1982) and Bollerslev (1986) explicitly rely on constancy of the unconditional second moments. In addition, existence of fourth moments is usually assumed in ARCH and GARCH models in order to derive simple estimators of the conditional variance parameters. These models have recently been extended to allow for drift in the unconditional variance (Engle and Bollerslev (1986)), leading to a long–run infinite unconditional variance; this extension raises the question whether stock returns, empirically, possess a unit root in variance. Hamilton (1988, 1989) formulates economic models with "switching regimes", in which each economic regime possesses a different (conditional) variance. After integrating the variance process over all possible regimes, one finds that this class of models also relies on constant unconditional variance. Given the role of covariance stationarity in all of this research, it would seem that testing for covariance stationarity and estimating the maximal moment exponent of stock returns series are both highly relevant to the debate on how to model the volatility observed in stock market returns data.

In this section, we analyze the behavior of the unconditional variance of two series of stock market returns: the first is a series of *monthly* returns on an aggregate stock portfolio from 1834 to 1987 $(n = 1848)$, which was also analyzed by Pagan and Schwert (1989a,b); the second is a series of *daily* returns to the "Standard & Poors 500" stock market index from 1962 to 1987 $(n = 6409)$, which was obtained from the 1988 CRSP tape. We are interested in the following questions: (i) what are the point estimates of the maximal moment exponent α of the two empirical time series, and is there evidence of fourth moment or second moment condition failure in the data, (ii) does the hypothesis of Pareto–like tails fit the data well, with a tail parameter α which is the same for both tails of the distribution and which is constant over time, and (iii) can we support the finding of Pagan and Schwert (1989a), who strongly rejected the null hypothesis of covariance stationarity for their series of monthly returns, when we perform tests that use the modified critical values which apply when fourth moments are not finite?

We chose to work with two long time series of stock returns since we are interested in incorporating information about "outlier" activity into our tests for covariance stationarity; smaller samples would provide us with fewer observations of "outliers", so that we would be unable to estimate the parameters of the tails of the distributions with enough precision. We will explicitly test for parameter nonconstancy, the two parameters of interest being the maximal moment exponent α and the unconditional variance. (When $\alpha < 2$, we would consider the scale parameter d rather than the variance.) An important advantage of studying data from financial markets is that they are reported "as is," without extreme observations having been smoothed over. Statistical smoothing procedures are typically applied to the raw data reported to the Commerce Department in order to generate time series of aggregate flow series such as consumption, investment, and national income; there is reason to believe that such smoothing, while appropriate for some purposes, may "remove" outliers and thus understate the leptokurtosis of aggregate flow data.

In estimating the variance of a distribution and testing for its constancy over time, it is important to obtain a preliminary estimate of the conditional mean of the series. Both returns series exhibit some evidence of calendar (weekday and month–of–year) effects. Therefore, the monthly series were regressed on 12 monthly dummies, and the daily series were regressed on 5 weekday as well as (in a second regression) on 12 monthly dummy variables; we retrieved the residuals from these regressions. The residual series further displayed a small but significant dependence on their own past values. Serial dependence in the presence of heavy tails of the marginal distributions prevents the cusum of squares statistic $\psi_n(r)$ from converging weakly in the Skorohod J_1 metric in $D[0,1]$, as we observed in Section 3.2. If the serial dependence is generated by a linear process as in (2), regressing out a (possibly long) AR process from the series yields asymptotically uncorrelated residuals, so that the statistic $\psi_n^\epsilon(r)$ converges weakly to $L_{\alpha/2}(r)$. For our data sets, regressing out an AR(12) process from the monthly series and an AR(5) from the daily series proved to be sufficient to create serially uncorrelated residual series. All empirical findings reported below are for these transformed series. We also tested for and found evidence of ARCH effects in the data; we did not "remove" them since they do not affect the unconditional variance, which it is our purpose to analyze here. The data and the preliminary transformations required to make the data amenable to formal testing in our framework are described in greater detail in the Data Appendix. Our findings, incidentally, are not affected significantly by these transformations, presumably because of the quantitatively small magnitude of the calendar and time series effects present in the original series.

We now study the tail behavior of the two series more closely, and in particular estimate the maximal moment exponent α of the distributions. Figure 9a graphs the right tail of the empirical cumulative distribution function (cdf) of the monthly series in double-logarithmic coordinates. (Formally, we graph $\log_{10}x$ against $P(X > \log_{10}x)$, for $x > 0$.) In these coordinates, Pareto–like tails of the distributions form straight lines with slope equal to $-\alpha$. As can be seen from the figure, about 5 to 10% of the observations fall into the right tail of the distribution. Figure 9b graphs the right tail of the cdf of the daily returns series in double–logarithmic coordinates; here too about 5 to 10% of the data fall into the right tail which seems to be very well characterized by a Pareto law. Similar findings obtain for the left tails (not shown here) of the two cdf's. We formally estimate the maximal moment exponent using formula (23) above. We computed α for a variety of choices of s, the number of included order statistics. We chose values of s such that we would not use observations that do not belong to the tails of the distributions. In the context of robust estimation of stable law parameters, Dumouchel (1983) has suggested that, as a rule of thumb, no more than 10% of a sample fall into the right or left tail of the distribution; this guideline seems to apply to our two data sets as well. This suggests that a conservative upper bound for a choice of s would be about 100 for the monthly series and about 250 for the daily series. (Point estimates of α decline rapidly for choices of s above these thresholds.) We report point estimates of α for several choices of s in Tables 5a and 5b for the monthly and daily series, respectively. The associated (asymptotic) standard errors are delivered from equation (25) above, and were computed under the assumption that $s = o(n^{2\gamma/(2\gamma+\alpha)})$, for some $\gamma > 0$. (Note that the standard errors could be sharpened considerably if we assumed that the tails were exactly Pareto, setting $\gamma = \infty$ and thus $s = o(n)$.) The point estimates are all below 4; they range from about 2.5 to 3.2 for the monthly series and from 3.1 to 3.8 for the daily series. The estimates exhibit some dependence on the choice of s, but it is important to note that they are almost all (at least for the larger values of s) more than two asymptotic standard deviations away from 4, so that we can be confident in concluding that the standard fourth moment condition is not met by either series.

Our point estimates of α are also all above 2, implying that the unconditional second moments are finite. While the tails of the two empirical distributions are heavier than those of the normal distribution, they do not seem to be heavy enough to fall into the domain of attraction of a stable distribution with $\alpha < 2$. Our direct estimates of the maximal moment exponent α of the distributions contribute a new element to the long-standing debate on whether to model stock returns in terms of stable laws. Mandelbrot's (1963) seminal work studied the behavior of relative price fluctuations of commodities such as cotton; he considered several pieces of evidence, among them recursive variance plots and graphs of the tails of the distributions in double–logarithmic coordinates, and argued that these were all strongly suggestive of the stable law behavior. Subsequent work by other researchers has generally concentrated on stock returns and foreign exchange rate data (e.g., Fama (1965), Blattberg and Gonedes (1974), Fielitz and Rozelle (1982), and more recently Akgiray and Booth (1988) and Hall, Brorsen and Irwin (1989)). The general conclusion to emerge from this literature is that empirical distributions in economics, especially aggregate series such as stock market returns, do not follow stable laws and are better modeled by finite variance distributions. Our point estimates of α agree with this general result. But the observation that variances are finite obviously does not suffice as a characterization of tail behavior. In particular it does not tell us the order of magnitude of the tail or how many moments of the distribution can be assumed to be finite. Since our point estimates of α are significantly below 4, fourth moment condition failure is a persistent feature of the data and affects the way tests for covariance stationarity must be carried out.

We recognize that empirical distributions have finite support and finite moments of all orders. It might therefore be argued that moment condition failure is merely an artificial by–product of the choice of distributional framework, here the model of Paretian tails. But such an argument would also preclude the use of *any* distribution that has infinite support, including the normal distribution, and does not provide a framework to describe the "outlier" activity present in the data. The "outliers" in our two series, as evidenced by the form of the cdf's plotted in Figures 9a and 9b, are rather well described as being distributed according to a Pareto law. Note that these plots too (by extrapolation) would assign a negligible probability of observing daily stock market returns of plus or minus 100% per day, say. Therefore, observing that the support of empirical distributions is bounded may not say much at all about the type of outlier activity that occurs in the data. In contrast, a model of Paretian tails not only appears to provide an adequate fit to observed outlier activity, but gives a predictive framework for the rate at which outliers appear and finally permits the development of an asymptotic distribution theory for tests of covariance stationarity when "outlier activity" plays a significant role.

To assess the robustness of our finding that the tails of stock returns distributions are Pareto–like, we formally tested the equality of estimates of α across the right and left tail and across time periods, using the sample split prediction test statistic $\tilde{V}(\tau_\alpha)$ proposed in section 3.6 above; the results are reported in Tables 6a and 6b. From (30), we may use standard normal critical values to evaluate the test results. We find that we cannot reject the null hypothesis that the estimates of α are the same for both tails, for either empirical distribution, and for all of our choices of s. Similarly, applying the $\tilde{V}(\tau_\alpha)$ test to two subsamples of the data equally split over time, we cannot reject the null that α is constant across time periods as well. These findings are important for our methodology, since we are interested in testing for a change of dispersion (or variance) over time, while taking α to be a constant parameter in our maintained hypothesis (C1).

We now formally test for covariance stationarity of the data. To simplify the discussion, we shall set $\hat{\alpha} = 3$ for the monthly series and $\hat{\alpha} = 3.5$ for the daily series. (None of our conclusions are affected by this simplification.) To estimate v, the lag truncation number ℓ was set, somewhat arbitrarily, equal to 8 for the monthly series and to 12 for the daily series. As in Pagan and Schwert (1989a), we set $k = 1$ for the sample split prediction test $V(\tau)$, and obtained test statistics of –3.34 for the monthly returns series (against a 2–sided 99% critical value of –2.34, for $\alpha = 3$), and –2.06 for the daily returns series (the 2–sided 95% critical value is –1.91 for $\alpha = 3.5$). We therefore reject the null of constant unconditional variance for both series, more strongly so for the monthly series. This result conforms with our conjecture that changes in the unconditional variance are a natural consequence of institutional change: since the monthly returns series covers 154 years vs. 25½ years for the daily returns series, it would be subject to more changes in stock market institutions and regulatory arrangements and thus may be expected to exhibit stronger evidence of failure of covariance stationarity.

The cusum of squares test $\psi_n^\epsilon(r)$ leads to the same conclusion with respect to covariance stationarity. The empirical cusum of squares statistics for the two series are graphed in Figures 10a and 10b. The minima of the two cusums are −2.13 and −1.61, respectively; the applicable critical values are given in Table 3, they are, for a 99% confidence level, equal to −1.40 $(\alpha = 3)$ and −1.44 $(\alpha = 3.5)$. We thus strongly reject the null of covariance stationarity for both series based on the test $\min_r(\psi_n^\epsilon(r))$. The range statistic $R_L = \max_r(\psi_n^\epsilon(r)) - \min_r(\psi_n^\epsilon(r)$ of the cusums are equal to 2.36 and 1.62, respectively; compare these values to a 99% critical value of 1.80 $(\alpha = 3)$ and a 90% critical value of 1.51 $(\alpha = 3.5)$. Based on the range statistics, therefore, we again reject the null of constant unconditional variance for both series, and again more strongly so for the monthly returns series. These results are all the more notable since the tests for covariance stationarity, as shown in Section 3, have low power against the alternative hypothesis of changing variance in the presence of fourth moment condition failure. We also note that both cusum of squares tests lie outside the respective 99% fdd critical value contours for a large range of choices of r, further strengthening our conclusion regarding covariance stationarity of the series.

Our empirical findings are thus twofold. First, both series considered here fail all of our tests for covariance stationarity. This throws into question the validity and robustness of the many studies that routinely employ this assumption while analyzing stock market volatility. Second, heavy tails are prominent features of the data and these tails are well described as being of the Pareto form with a constant tail exponent α. In the absence of formal economic models that provide plausible mechanisms for generating such heavy-tailed series, we do not have a framework to assess the theory content of this empirical finding. However, the apparent stability of this tail shape parameter over very long periods of time is an interesting empirical regularity that models of rational economic behavior should be designed to accommodate and explain.

6. CONCLUSION

This paper develops a limit theory for tests of covariance stationarity in the presence of heavy tailed distributions. Sample split prediction tests and studentized cusum of squares tests are based on estimates of second and fourth moments of the data. When the usual fourth moment condition holds, standard normal and Brownian bridge asymptotics apply. When fourth moments are infinite, the limit laws of these tests are given by functionals of stable processes. Both these tests for covariance stationarity are consistent as long as second moments are finite, but they are conservative when evaluated at conventional critical values, and they have low power against the alternative of unconditional heteroskedasticity. When second moments are infinite as well, the tests are inconsistent.

The paper suggests a new test for heterogeneity which is robust to moment condition failure. Instead of relying on estimated moments, the test is based on direct estimation of the tail parameters of a family of distributions with Pareto–like tails. The test suggested has a standard normal limiting distribution. The "cost" of its robustness is that it has lower power than moments–based tests when moment condition failure is not at issue. Based on Monte Carlo simulations, we provide revised critical values for the sample split prediction test and the studentized cusum of squares test. These new critical values should be used when fourth moments are not finite.

In an empirical application, we test whether monthly and daily stock market returns are covariance stationary. We find that both series are characterized by heavy tails and fourth moment condition failure. Neither series passes the tests for covariance stationarity, an empirical finding that confirms earlier work of Pagan and Schwert (1989a) and casts doubts on the validity and descriptive accuracy of econometric models that assume the unconditional variance of stock market returns to be constant. The maximal moment exponent α of the distributions is found to be constant over time for both series, indicating that even when fourth moment conditions and covariance stationarity fail there are still interesting empirical regularities in the series.

We close by mentioning some possible extensions to our work. On the analytical side, we may go beyond the class of linear models whose independent innovations have Paretian tails, and study models of conditional heterogeneity, mixture distributions and the ways in which these models affect tests for covariance stationarity. To obtain the small sample properties of our tests, we would need to strengthen the distributional assumptions made in our paper and assume, say, exactly Pareto–distributed or t–distributed innovations. On the empirical side, a vast amount of work beckons: how common is unconditional heteroskedasticity in economic time series, and how common is moment condition failure? Do individual returns series behave similarly to the aggregate stock market index series we have considered here? Do Paretian tails characterize economic and financial time series other than the stock market series? Finally, what theoretical models of rational economic behavior would plausibly explain and predict the apparent constancy of the nature of the outlier activity of the data? All of these issues seem worthy of future research.

MATHEMATICAL APPENDIX

1. PROOF OF LEMMA 2.1

(a) Set $x_t' = (y_{t-1}, \ldots, y_{t-p})$ and write $\hat{\epsilon}_t = \epsilon_t + (\varphi-\hat{\varphi})'x_t$. Then

$$\text{(A1)}\quad \hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2 = \epsilon_t^2 - \sigma_\epsilon^2 + (\sigma_\epsilon^2 - \hat{\sigma}_\epsilon^2) + 2(\varphi-\hat{\varphi})'x_t\epsilon_t + (\varphi-\hat{\varphi})'x_tx_t'(\varphi-\hat{\varphi}) .$$

$$\begin{aligned}\text{(A2)}\quad (nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) &= (nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_\epsilon^2) - ([nr]/n)(nv_\epsilon)^{-1/2}\Sigma_1^n(\epsilon_t^2 - \sigma_\epsilon^2)\\ &+ 2(\varphi-\hat{\varphi})'[(nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}x_t\epsilon_t + (\varphi-\hat{\varphi})'[(nv)^{-1/2}\Sigma_1^{[nr]}x_tx_t'](\varphi-\hat{\varphi})\\ &- ([nr]/n)\{2(\varphi-\hat{\varphi})'[(nv_\epsilon)^{-1/2}\Sigma_1^n x_t\epsilon_t] + (\varphi-\hat{\varphi})'[(nv)^{-1/2}\Sigma_1^n x_tx_t'](\varphi-\hat{\varphi})\}\end{aligned}$$

and by Donsker's theorem for partial sums of iid variates (e.g. Billingsley (1968), p. 137) it follows that

$$\text{(A3)}\quad (nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_\epsilon^2) - ([nr]/n)(nv_\epsilon)^{-1/2}\Sigma_1^n(\epsilon_t^2 - \sigma_\epsilon^2) \to_d W(r) - rW(1) = B(r) .$$

Thus, we deduce

$$\text{(A4)}\qquad (nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) \to_d B(r)$$

from (A2) and (A3) provided

$$\text{(A5)}\qquad \sup_r|2(\varphi-\hat{\varphi})'(nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}x_t\epsilon_t + (\varphi-\hat{\varphi})'[(nv)^{-1/2}\Sigma_1^{[nr]}x_tx_t'](\varphi-\hat{\varphi})| \to_p 0 .$$

This follows if

$$\text{(A6)}\qquad (\varphi-\hat{\varphi})'\max_k[(nv_\epsilon)^{-1/2}\Sigma_1^k x_t\epsilon_t] \to_p 0$$

$$\text{(A7)}\qquad (\varphi-\hat{\varphi})'\max_k[(nv_\epsilon)^{-1/2}\Sigma_1^k x_tx_t'](\varphi-\hat{\varphi})' \to_p 0 .$$

But $\Sigma_1^k x_t\epsilon_t$ is a martingale and by the martingale maximal inequality (e.g. Hall and Heyde (1980), p. 14)

$$P\left[\max_k|n^{-3/4}\Sigma_1^k x_t\epsilon_t| > \delta\right] < \delta^{-2}n^{-3/2}\Sigma_1^n\sigma_\epsilon^2E(x_t'x_t) = p\delta^{-2}n^{-1/2}\sigma_\epsilon^2\sigma_y^2 \to 0$$

for all $\delta > 0$. Thus,

$$(\varphi-\hat{\varphi})'\max_k[(nv_\epsilon)^{-1/2}\Sigma_1^k x_t\epsilon_t] = n^{1/4}(\varphi-\hat{\varphi})'\max_k[n^{-3/4}v_\epsilon^{-1/2}\Sigma_1^k x_t\epsilon_t] \to_p 0 ,$$

since $n^{1/2}(\hat{\varphi}-\varphi) = O_p(1)$, and (A6) holds. (A7) is immediate since the left side is dominated by

$$v_\epsilon^{-1/2}(\varphi-\hat{\varphi})'[n^{-1}\Sigma_1^n x_t x_t']n^{1/2}(\varphi-\hat{\varphi})$$

which tends to zero in probability. Hence (A5) holds and we have (A4) as required.

(b) We consider three cases.

Case (i). $2 < \alpha < 4$. Here $\sigma_\epsilon^2 = E(\epsilon_t^2) < \infty$ and we have a decomposition that is analogous to (A2), viz.

(A8)
$$\begin{aligned} a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) &= a_n^{-2}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_\epsilon^2) - ([nr]/n)a_n^{-2}\Sigma_1^n(\epsilon_t^2 - \sigma_\epsilon^2) \\ &+ 2(\varphi-\hat{\varphi})'[a_n^{-2}\Sigma_1^{[nr]}x_t\epsilon_t] + (\varphi-\hat{\varphi})'[a_n^{-2}\Sigma_1^{[nr]}x_t x_t'](\varphi-\hat{\varphi}) \\ &- ([nr]/n)\{2(\varphi-\hat{\varphi})'[a_n^{-2}\Sigma_1^n x_t x_t'] + (\varphi-\hat{\varphi})'[a_n^{-2}\Sigma_1^n x_t x_t'(\varphi-\hat{\varphi})\} . \end{aligned}$$

Now

(A9)
$$a_n^{-2}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_\epsilon^2) \to_d U_{\alpha/2}(r)$$

by (6), since $\epsilon_t^2 \in N\mathcal{D}(\alpha/2)$ and $\epsilon_t^2 - \sigma_\epsilon^2$ is an iid sequence with zero mean. It follows from (A8) and (A9) that

(A10)
$$a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) \to_d U_{\alpha/2}(r) - rU_{\alpha/2}(1)$$

provided

(A11)
$$\sup_r|2(\varphi-\hat{\varphi})'(a_n^{-2}\Sigma_1^{[nr]}x_t\epsilon_t) + (\varphi-\hat{\varphi})'(a_n^{-2}\Sigma_1^{[nr]}x_t x_t')(\varphi-\hat{\varphi})| \to_p 0 ,$$

since (A11) implies the corresponding Skorohod distance necessarily converges to zero in probability. Thus, (A10) converges weakly to the limit process $U_{\alpha/2}(r)$ in $D[0,1]$ in the J_1–Skorohod topology if (A11) holds. But (A11) follows if

(A12)
$$(\varphi-\hat{\varphi})'\max_k|a_n^{-2}\Sigma_1^k x_t\epsilon_t| \to_p 0 ,$$

since the contribution from the second term inside the sup is dominated by a quantity that tends to zero in probability because $\hat{\varphi}-\varphi \to_p 0$, viz.

$$(\varphi-\hat{\varphi})'\max_k[a_n^{-2}\Sigma_1^k x_t x_t'](\varphi-\hat{\varphi})$$
$$\leq (\varphi-\hat{\varphi})'[a_n^{-2}\Sigma_1^n x_t x_t'](\varphi-\hat{\varphi}) \to_p 0 .$$

However,

$$P\left[\max_k |a_n^{-2}\Sigma_1^k x_t \epsilon_t| > \delta\right] < \delta^{-2} a_n^{-4}\Sigma_1^n \sigma_\epsilon^2 E(x_t' x_t)) = p\delta^{-2}\sigma_\epsilon^2\sigma_y^2(n/a_n^4) \to 0$$

and (A12) holds, giving (A10) as required.

Case (ii). $\alpha = 2$. In this case we set $\sigma_{\epsilon n}^2 = E[\epsilon_t^2 1(\epsilon_t^2 < a_n^2)]$. The decomposition (A8) has the same form as before with $\sigma_{\epsilon n}^2$ replacing σ_ϵ^2. In place of (A9) we have

$$a_n^{-2}\Sigma_1^k(\epsilon_t^2 - \sigma_n^2) \to_d U_{\alpha/2}(r)$$

and (A10) follows since (A11) holds as in case (i).

Case (iii). $0 < \alpha < 2$. In this case, no centering is required and the first two terms on the right side of (A8) are replaced by

$$a_n^{-2}\Sigma_1^k\epsilon_t^2 - ([nr]/n)a_n^{-2}\Sigma_1^n\epsilon_t^2 \to_d U_{\alpha/2}(r) - rU_{\alpha/2}(1) .$$

Here, the limit process $U_{\alpha/2}(r)$ is a positive stable process on $D[0,1]$. As in Case (i), (A11) holds if (A12) does. But

$$\max_k |a_n^{-2}\Sigma_1^k x_t\epsilon_t| \leq a_n^{-2}\Sigma_1^n |x_t\epsilon_t| \leq \left[a_n^{-2}\Sigma_1^n x_t^2\right]^{-1/2}\left[a_n^{-2}\Sigma_1^n\epsilon_t^2\right]^{1/2} = O_p(1)$$

since $\epsilon_t^2, x_t^2 \in \mathcal{D}(\alpha/2)$. Finally, by Hannan and Kanter (1977)), $n^{1/\delta}(\hat{\varphi}-\varphi) \to_{a.s.} 0$ for any $\delta > \alpha$, so that $\hat{\varphi}-\varphi \to_{a.s.} 0$ as (A12) holds, giving the required result (A10) with $U_{\alpha/2}(r)$ being a positive stable process in this case.

2. PROOF OF THEOREM 3.1

(a) From Theorem 3.7 of Phillips and Solo (1989) we have the following CLT for variances under (C1)–(C3) and $\alpha > 4$:

(A13) $$n^{-1/2}\Sigma_1^n(y_t^2 - \sigma_y^2) \to_d N(0,v)$$

where $\sigma_y^2 = E(y_t^2) = \sigma_\epsilon^2 \Sigma_0^\infty c_j^2$ and

(A14) $$v = \kappa_4 \left[\Sigma_0^\infty c_j^2\right]^2 + 2\sigma_\epsilon^4 \Sigma_{-\infty}^\infty \left[\Sigma_{s=0}^\infty c_s c_{s+j}\right]^2 .$$

Note that κ_4 is the fourth cumulant of ϵ_t and the series that appear in v converge under (C3) by Lemma 3.5 of Phillips and Solo (1989). Next, observe that

$$n_1^{1/2}\hat{\tau} = n_1^{-1/2}\Sigma_1^{n_1}(y_t^2 - \sigma_y^2) - (n_1/n_2)^{1/2} n^{-1/2} \Sigma_{n_1+1}^n (y_t^2 - \sigma_y^2)$$

and, by assumption, $k_n = n_1/n_2 \to k$ as $n \to \infty$. The two sums that appear in this expression are asymptotically independent under (C3) and again by Theorem 3.7 of Phillips and Solo (1989) we deduce that

$$n_1^{1/2}\tau \to_d N(0, (1+k)v) .$$

Finally, $\hat{v} \to_p v$ since $\hat{v}$ is the usual Bartlett estimate of the long run variance of y_t^2 and this estimate is consistent under the stated conditions. The required result, viz. $V_k(\tau) \to_d N(0,1)$, now follows directly.

(b) Suppose $2 < \alpha < 4$, so that $E(y_t^2) = \sigma_y^2 = (\Sigma_0^\infty c_j^2)\sigma_\epsilon^2 < \infty$. Set $a_n = an^{1/\alpha}$, $r = k/(1+k)$ and note that

$$\frac{n_1}{n} = \frac{k_n}{1 + k_n} = \frac{[nr]}{n} + o(1) \to r ,$$

$$\frac{n_2}{n} = \frac{1}{1 + k_n} = \frac{n-[nr]}{n} + o(1) \to 1-r .$$

Next, observe that

$$y_t^2 = \Sigma_0^\infty c_j^2 \epsilon_{t-j}^2 + \Sigma_{j=0}^\infty \Sigma_{\substack{r=-j \\ r \neq 0}}^\infty c_j c_{j+r} \epsilon_{t-j} \epsilon_{t-j-r}$$

(A15) $$= f_0(L)\epsilon_t^2 + \Sigma_{r=-\infty,\neq 0}^\infty f_r(L)\epsilon_t \epsilon_{t-r}$$

where $f_h(L) = \Sigma_0^\infty c_j c_{j+h}$. We use the following decomposition of the lag polynomial $f_h(L)$ (see equation (23) and Lemma 2.1 of Phillips and Solo (1989)):

(A16) $$f_h(L) = f_h(1) - (1-L)\tilde{f}_h(L)$$

where

$$\tilde{f}_h(L) = \Sigma_{k=0}^{\infty}\tilde{f}_{hk}L^k\ ,\ \ \tilde{f}_{hk} = \Sigma_{s=k+1}^{\infty}f_{hs} = \Sigma_{s=k+1}^{\infty}c_s c_{s+h}\ .$$

Employ (A16) in (A15) and note that $\epsilon_t^2 \in \mathcal{ND}(\alpha/2)$ so that $a_n^{-2}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_\epsilon^2) \rightarrow_d U_{\alpha/2}(r)$ by (5); while for $s \neq 0$ $\epsilon_t\epsilon_{t-s} \in \mathcal{ND}(2)$ and we have $a_n^{-2}\Sigma_1^{n_1}\epsilon_t\epsilon_{t-s} = o_p(1)$. Combining these results we obtain

$$\begin{aligned} a_n^{-2}\Sigma_1^{n_1}(y_t^2 - \sigma_y^2) &= a_n^{-2}\Sigma_1^{[nr]}(y_t^2 - \sigma_y^2) + o_p(1) \\ &= f_0(1)[a_n^{-2}\Sigma_1^{[nr]}(\epsilon_t^2 - \sigma_t^2)] + o_p(1) \\ &\rightarrow_d U_{\alpha/2}(r)\ , \end{aligned} \tag{A17}$$

with $r = k/(1+k)$ fixed and

$$\sigma^2 = f_0(1) = \Sigma_0^{\infty}c_j^2\ .$$

In a similar way we deduce

(A18) $$a_n^{-2}\Sigma_{n_1+1}^{n}(y_t^2 - \sigma_y^2) \rightarrow_d \sigma^2[U_{\alpha/2}(1) - U_{\alpha/2}(r)]\ .$$

It follows from (A17) and (A18) that

$$\begin{aligned} n_1 a_n^{-1}\hat{\tau} &= a_n^{-2}\Sigma_1^{n_1}(y_t^2 - \sigma_y^2) - k_n a_n^{-2}\Sigma_{n_1+1}^{n}(y_t^2 - \sigma_y^2) \\ &\rightarrow_d \sigma^2\{U_{\alpha/2}(r) - k[U_{\alpha/2}(1) - U_{\alpha/2}(r)] \\ &= \sigma^2\{(1+k)U_{\alpha/2}(k/(1+k)) - kU_{\alpha/2}(1)\}\ . \end{aligned} \tag{A19}$$

Turning to the denominator of $V_k(\tau)$ and using the same arguments as those in Phillips (1990, Section 2.3) we find

(A20) $$na_n^{-4}\hat{v} \rightarrow_d \sigma^4\int_0^1(dU_{\alpha/2})^2\ ,$$

so that

(A21) $$n_1(1+k_n)a_n^{-4}\hat{v} = k_n(na_n^{-4}\hat{v}) \to_d k\sigma^4 \int_0^1 (dU_{\alpha/2})^2 .$$

It follows that

$$V_k(\tau) = [(1+k_n)\hat{v}]^{-1/2}(n_1^{1/2}\tau) = \left[n_1(1+k_n)a_n^{-4}\hat{v}\right]^{-1/2}(n_1 a_n^{-2}\hat{\tau})$$
$$\to_d \left[k\int_0^1 (dU_{\alpha/2})^2\right]^{-1/2}[(1+k)U_{\alpha/2}(k/(1+k)) - kU_{\alpha/2}(1)]$$

as required.

When $\alpha = 2$ we may center on $\sigma_{yn}^2 = E[y_t^2 1(y_t^2 < a_n^2)]$, giving

$$n_1 a_n^{-2}\hat{\tau} = a_n^{-2}\Sigma_1^{n_1}(y_t^2 - \sigma_{yn}^2) - k_n a_n^{-2}\Sigma_{n_1+1}^{n}(y_t^2 - \sigma_{yn}^2)$$
$$\to_d \sigma^2\{U_{\alpha/2}(r) - k[U_{\alpha/2}(1) - U_{\alpha/2}(r)]\}$$

as in (A19). (A21) continues to hold when $\alpha = 2$, and this covers the case $\alpha = 2$.

When $0 < \alpha < 2$ we note that $\alpha/2 < 1$ and so no centering is needed in the numerator of the statistic. In this case we have

$$n_1 a_n^{-2}\hat{\tau} = a_n^{-2}\Sigma_1^{n_1} y_t^2 - k_n a_n^{-2}\Sigma_{n_1+1}^{n} y_t^2$$
$$\to_d \sigma^2[U_{\alpha/2}^+(r) - k\{U_{\alpha/2}^+(1) - U_{\alpha/2}^+(r)\}]$$

where $U_{\alpha/2}^+(r)$ is a positive stable process with exponent $\alpha/2$ i.e. increments in $U_{\alpha/2}^+(r)$ are independent and are distributed as a positive stable variate with characteristic exponent $\alpha/2$. (A21) holds as before and this establishes part (b). □

3. PROOF OF COROLLARY 3.2

We need to show that the limit variate given by (11) is equivalent in distribution to (12). First observe that the variates $U_{\alpha/2}(r) - U_{\alpha/2}(r/2)$ and $U_{\alpha/2}(r/2)$ are stable, independent and equivalent in distribution, i.e.

$$U_{\alpha/2}(r) - U_{\alpha/2}(r/2) =_d U_{\alpha/2}(r/2) .$$

Hence

$$U_{\alpha/2}(r/2) - \{U_{\alpha/2}(r) - U_{\alpha/2}(r/2)\} =_d -[U_{\alpha/2}(r/2) - \{U_{\alpha/2}(r) - U_{\alpha/2}(r/2)\}]$$

and

$$U^s_{\alpha/2}(r) = U_{\alpha/2}(r/2) - \{U_{\alpha/2}(r) - U_{\alpha/2}(r/2)\}$$

is a symmetric stable process on $D[0,1]$. It follows that

$$U_{\alpha/2}(1/2) - U_{\alpha/2}(1) = U_{\alpha/2}(1/2) - \{U_{\alpha/2}(r) - U_{\alpha/2}(r/2)\} = U^s_{\alpha/2}(1)$$

as required for the numerator of (12).

Now let $V_{\alpha/2}(r/2) = U_{\alpha/2}(r) - U_{\alpha/2}(r/2)$ and write

$$U^s_{\alpha/2}(r) = U_{\alpha/2}(r/2) - V_{\alpha/2}(r/2)$$

as the difference of the two independent stable processes $U_{\alpha/2}$ and $V_{\alpha/2}$. Increments in these processes are also independent and we have

$$dU^s_{\alpha/2}(r) = dU_{\alpha/2}(r/2) - dV_{\alpha/2}(r/2) =_d (1/2)^{2/\alpha} dU_{\alpha/2}(r) - (1/2)^{2/\alpha} dV_{\alpha/2}(r)\,.$$

Next observe that

$$dU_{\alpha/2}(r),\ dV_{\alpha/2}(r) \in \mathcal{ND}(\alpha/2)$$

and hence

$$(dU_{\alpha/2}(r))^2,\ (dV_{\alpha/2}(r))^2 \in \mathcal{ND}(\alpha/4)\,.$$

However, in view of the independence of $U_{\alpha/2}(r)$ and $V_{\alpha/2}(r)$ products of increments in these processes are in $\mathcal{D}(\alpha/2)$ and hence

$$dU_{\alpha/2}(r)dV_{\alpha/2}(r) = 0 \text{ a.s.}$$

Thus

$$\begin{aligned}(dU^s_{\alpha/2}(r))^2 &=_d (1/2)^{4/\alpha}(dU_{\alpha/2}(r))^2 + (1/2)^{4/\alpha}(dV_{\alpha/2}(r))^2 \\ &=_d [(1/2) + (1/2)]^{4/\alpha}(dU_{\alpha/2}(r))^2 \\ &= (dU_{\alpha/2}(r))^2\,.\end{aligned}$$

It follows that

$$\left[\int_0^1 (dU_{\alpha/2})^2\right]^{-1/2} [2U_{\alpha/2}(1/2) - U_{\alpha/2}(1)] =_d \left[\int_0^1 (dU^s_{\alpha/2})^2\right]^{-1/2} U^s_{\alpha/2}(1)$$

thereby establishing (12) as required.

4. PROOF OF THEOREM 3.3

(a) When $\alpha > 4$, we have $\hat{v}_1, \hat{v}_2 \to_p v$ and $\hat{v}_1 + k_n \hat{v}_2 \to_p (1+k)v$. Further, $n_1^{1/2}\hat{\tau} \to_d N(0, (1+k)v)$ as before and part (a) follows directly.

(b) When $0 < \alpha < 4$ the numerator of

$$\tilde{V}_k(\tau) = n_1^{1/2}\hat{\tau}/\tilde{v}^{1/2} = (n_1 a_n^{-2}\hat{\tau})/(n_1 a_n^{-4}\tilde{v})^{1/2}$$

has the same limit behavior as before, viz.

$$n_1 a_n^{-2}\hat{\tau} \to_d \sigma^2[U_{\alpha/2}(r) - k\{U_{\alpha/2}(1) - U_{\alpha/2}(r)\}] \tag{A22}$$

where $r = k/(1+k)$. The denominator behaves as follows:

$$n_1 a_n^{-4}\tilde{v} = n_1 a_n^{-4}\hat{v}_1 + k_n^2 (a_n^{-4} n_2 \hat{v}_2)$$

$$\to_d \sigma^4\left\{\int_0^r (dU_{\alpha/2})^2 + k^2 \int_0^1 (dU_{\alpha 2})^2\right\}. \tag{A23}$$

Combining (A22) and (A23) we get the first expression given in part (b). The distributional equivalence of (13) and (14) when $k = 1$ follows just as in Corollary 3.2.

5. PROOF OF THEOREM 3.4

In all cases (i.e. $0 < \alpha < 4$ and $\alpha \geq 4$) we have $\hat{\varphi} \to_p \varphi$. Next, when $\alpha > 2$ we have $\hat{\sigma}^2_{1\epsilon}, \hat{\sigma}^2_{2\epsilon} \to_p \sigma^2_\epsilon$; and when $\alpha > 4$ we have:

$$\hat{v}_{1\epsilon} = n_1^{-1}\Sigma_1^{n_1}\left[\hat{\epsilon}_t^2 - \hat{\sigma}^2_{1\epsilon}\right]^2 = n_1^{-1}\Sigma_1^{n_1}\left[\epsilon_t^2 - \sigma_\epsilon^2\right]^2 + o_p(1)$$

$$\to_p v_\epsilon = \kappa_4 + 2\sigma_\epsilon^4$$

and, similarly, $\hat{v}_{2\epsilon} \to_p v_\epsilon$. Thus, when $\alpha > 4$, we have $\tilde{v} \to_p (1+k)v_\epsilon$ and

$$n_1^{1/2}\hat{\tau}_\epsilon = n_1^{-1/2}\Sigma_1^{n_1}(\hat{\epsilon}_t^2 - \hat{\sigma}_{1\epsilon}^2) - k_n^{1/2} n_2^{-1/2}\Sigma_{n_1+1}^{n}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2)$$
$$\to_d N(0, (1+k)v_\epsilon) .$$

It follows that when $\alpha > 4$ we have

$$\tilde{V}_k(\tau_\epsilon) \to_d N(0,1)$$

as required.

When $2 < \alpha < 4$ we employ Lemma 2.1, set $r = k/(1+k)$ and find that

$$\text{(A24)} \qquad n_1 a_n^{-2}\hat{\tau}_\epsilon = a_n^{-2}\Sigma_1^{n_1}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) - k_n a_n^{-2}\Sigma_{n_1+1}^{n}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2)$$
$$= a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) - k_n a_n^{-2}\Sigma_{[nr]+1}^{n}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) + o_p(1)$$
$$\text{(A25)} \qquad \to_d \sigma^2\{U_{\alpha/2}(r) - k[U_{\alpha/2}(1) - U_{\alpha/2}(r)]\}$$

and

$$n_1 a_n^{-4}\tilde{v}_\epsilon = n_1 a_n^{-4}\hat{v}_{1\epsilon} + k_n^2(a_n^{-4} n_2 \hat{v}_{2\epsilon})$$
$$\text{(A26)} \qquad \to_d \sigma^4\{\textstyle\int_0^r (dU_{\alpha/2})^2 + k^2 \int_r^1 (dU_{\alpha/2})^2\}$$

as in (A19) and (A23). When $\alpha = 2$ we center on $\sigma_{\epsilon n}^2 = E(\epsilon_1^2 1(\epsilon_t^2 < a_n))$ in the sums in (A24) and, when $0 < \alpha < 2$, no centering of the sums in (A24) is required. In both cases (A25) holds, as does (A26). Writing

$$\tilde{V}_k(\tau_\epsilon) = n_1 a_n^{-2}\tilde{\tau}_\epsilon / \left[n_1 a_n^{-4}\tilde{v}_\epsilon\right]^{1/2}$$

and using (A25), (A26) and joint weak convergence we obtain the required asymptotic equivalence of $\tilde{V}_k(\tau_\epsilon)$ and $\tilde{V}_k(\tau)$. Similar derivations yield the asymptotic equivalence of $V_k(\tau_\epsilon)$ and $V_k(\tau)$. □

6. PROOF OF THEOREM 3.5

(a) Under H_1 we have

(A27) $$\hat{\tau} = n_1^{-1}\Sigma_1^{n_1}(y_t^2 - \sigma_y^2) - n_2^{-1}\Sigma_{s=1}^{n_2}(y_{n_1+s}^2 - \sigma_{ys}^2) + \sigma_y^2 - n_2^{-1}\Sigma_{s=1}^{n_2}\sigma_{ys}^2 ,$$

so that

$$n_1^{1/2}\hat{\tau} = n_1^{-1/2}\Sigma_1^{n_1}(y_t^2 - \sigma_y^2) - k_n^{1/2}n_2^{-1/2}\Sigma_{s=1}^{n_2}(y_{n_1+s}^2 - \sigma_{ys}^2) + k_n^{1/2}n_2^{-1/2}\Sigma_1^{n_2}(\sigma_y^2 - \sigma_{ys}^2) .$$

Now $k_n \to k$ as $n \to \infty$ and

$$\begin{aligned} n_2^{-1/2}\Sigma_{s=1}^{n_2}(\sigma_y^2 - \sigma_{ys}^2) &= n_2^{-1/2}\Sigma_{s=1}^{n_2}\Sigma_{j=0}^{s-1}c_j^2(\sigma_\epsilon^2 - \sigma_{\epsilon+}^2) \\ &= (\sigma_\epsilon^2 - \sigma_{\epsilon+}^2)n_2^{-1/2}\Sigma_{j=0}^{n_2}(n_2 - j)c_j^2 \\ &= (\sigma_\epsilon^2 - \sigma_{\epsilon+}^2)n_2^{1/2}\Sigma_{j=0}^{n_2}(1 - j/n_2)c_j^2 \\ &= O(n^{1/2}) . \end{aligned}$$

Hence, $n_1^{1/2}\hat{\tau} = O_p(n^{1/2})$. The denominators of $V_k(\hat{\tau})$ and $\tilde{V}_k(\hat{\tau})$ depend on $\hat{v}$ and $\tilde{v}$ and

$$\hat{v} \to_p v = \lim_{n\to\infty}\left\{n^{-1}E\left[\Sigma_1^n(y_t^2 - E(y_t^2))\right]^2\right\} ,$$

$$\tilde{v} = \hat{v}_1 + k_n\hat{v}_2 \to_p \lim_{n\to\infty} n_1^{-1}E\left[\Sigma_1^{n_1}(y_t^2 - E(y_t^2))\right]^2$$

$$+ k \lim_{n\to\infty} n_2^{-1}E\left[\Sigma_{n_1+1}^{n}(y_t^2 - E(y_t^2))\right]^2 .$$

These limits exist under (C3) since fourth moments of ϵ_t are finite and bounded. It follows that $V_k(\tau), \tilde{V}_k(\tau) = O_p(n^{1/2})$ as required and these tests are consistent under H_1. In an entirely analogous way we find that $V_k(\tau_\epsilon)$, $\tilde{V}_k(\tau_\epsilon) = O_p(n^{1/2})$. This establishes the rates of divergence under H_1' given in (15) for $\alpha > 4$.

Expression (A27) continues to hold when $2 < \alpha < 4$ and we have

$$n_1 a_n^{-2}\hat{\tau} = a_n^{-2}\Sigma_1^{n_1}(y_t^2 - \sigma_y^2) - k_n a_n^{-2}\Sigma_{s=1}^{n_2}(y_{n_1+s}^2 - \sigma_{ys}^2) + k_n a_n^{-2}\Sigma_1^{n_2}(\sigma_y^2 - \sigma_{ys}^2) .$$

Now we find

$$a_n^{-2}\Sigma_1^{n_2}(\sigma_y^2 - \sigma_{ys}^2) = (\sigma_\epsilon^2 - \sigma_{\epsilon+}^2)(n_2 a_n^{-2})\Sigma_{j=0}^{n_2}(1 - j/n_2)c_j^2$$
$$= O(n^{1-2/\alpha}) .$$

But

(A28) $$V_k(\tau) = \left[n_1(1 + k_n)a_n^{-4}\hat{v}\right]^{-1/2}[n_1 a_n^{-2}\hat{\tau}]$$

and the denominator $n_1(1 + k_n)a_n^{-4}\hat{v} = O_p(1)$ as in the null case. Thus, $V_k(\tau) = O_p(n^{1-2/\alpha})$ as asserted in (16). The same rate of divergence under H_1 applies to $\tilde{V}_k(\tau)$, $V_k(\tau_\epsilon)$ and $\tilde{V}_k(\tau_\epsilon)$, thereby establishing part (a).

(b) Observe that both ϵ_t^2 and y_t^2 lie in $\mathcal{ND}(\alpha/2)$. When $0 < \alpha < 2$, neither ϵ_t^2 nor y_t^2 has finite mean and no centering is required for the limit theory. Further, for $r = k/(1+k)$ fixed, we have

$$n_1 a_n^{-2}\hat{\tau} = a_n^{-2}\Sigma_1^{n_1} y_t^2 - k_n a_n^{-2}\Sigma_{n_1+1}^{n} y_t^2$$

(A29) $$\to_d \sigma^2\{U_{\alpha/2}^+(r) - kf^2[U_{\alpha/2}^+(1) - U_{\alpha/2}^+(r)]\}$$

where $\sigma^2 = \Sigma_0^\infty c_j^2$, as before, and $f = a_+/a$. To verify (A29) we employ (A15) as in the proof of Theorem 3.1, giving

$$a_n^{-2}\Sigma_1^{n_1} y_t^2 = \sigma^2(a_n^{-2}\Sigma_1^{n}\epsilon_t^2) + o_p(1) \to_d \sigma^2 U_{\alpha/2}^+(r)$$

and

$$a_n^{-2}\Sigma_{n_1+1}^{n} y_t^2 = \sigma^2(a_{+,n}/a_n)^2(a_{+,n}^{-2}\Sigma_{n_1+1}^{n}\epsilon_t^2) + o_p(1) \to_d \sigma^2 f^2\{U_{\alpha/2}^+(1) - U_{\alpha/2}^+(r)\}$$

as required, where $a_{+,n} = a_+ n^{1/\alpha}$ is the normalizing constant for the upper half sequence $(\epsilon)_{n_1+1}^\infty$.

For the denominator of $V_k(\tau)$ in (A28) we find

$$
\begin{aligned}
na_n^{-4}\hat{v} &= \sigma^4 a_n^{-4}\Sigma_1^n \epsilon_t^2 + o_p(1) \\
&= \sigma^4\left\{a_n^{-4}\Sigma_1^{n_1}\epsilon_t^2 + (a_{+,n}/a_n)^4(a_{+,n}^{-4}\Sigma_{n_1+1}^n \epsilon_t^2)\right\} + o_p(1) \\
&\to_d \sigma^4\left\{\int_0^r\left[dU_{\alpha/2}^+\right]^2 + f^4\int_r^1(dU_{\alpha/2})^2\right\}.
\end{aligned}
$$

Hence

$$
n_1(1+k_n)a_n^{-4}\hat{v} \to k\sigma^4\left\{\int_0^r\left[dU_{\alpha/2}^+\right]^2 + f^4\int_r^1(dU_{\alpha/2})^2\right\}.
$$

We deduce that under H_1'

$$
V_k(\tau) \to_d \left[k\left\{\int_0^r\left[dU_{\alpha/2}^+\right]^2 + f^4\int_r^1\left[dU_{\alpha/2}^+\right]^2\right\}\right]^{-1/2}[U_{\alpha/2}^+(r) - kf^2\{U_{\alpha/2}^+(1) - U_{\alpha/2}(r)\}]
$$

and, hence, $V_k(\tau) = O_p(1)$, leading to an inconsistent test. Similar derivations show that $\tilde{V}_k(\tau)$, $V_k(\tau_\epsilon)$ and $\tilde{V}_k(\tau_\epsilon)$ are $O_p(1)$ as $n \to \infty$ and these tests are also inconsistent under H_1', thereby establishing part (b). □

7. PROOF OF THEOREM 3.6

(a) Write

$$
\begin{aligned}
\psi_n(r) &= (n\hat{v})^{-1/2}\Sigma_1^{[nr]}(y_t^2 - \sigma_y^2 + \sigma_y^2 - \hat{\mu}_2) \\
&= (n\hat{v})^{-1/2}\Sigma_1^{[nr]}(y_t^2 - \sigma_y^2) + [nr](n\hat{v})^{-1/2}(\sigma_y^2 - \hat{\mu}_2) \\
&= (v/\hat{v})^{1/2}\left[(nv)^{-1/2}\Sigma_1^{[nr]}(y_t^2 - \sigma_y^2) - ([nr]/n)(nv)^{-1/2}\Sigma_1^n(y_t^2 - \sigma_y^2)\right] \\
&\to_d W(r) - rW(1)
\end{aligned}
$$

by Theorem 3.7 of Phillips and Solo (1989) since $\hat{v}/v \to_p 1$. The limit process $B(r) = W(r) - rW(1)$ is a Brownian bridge on $C[0,1]$.

Similarly, we have

$$\psi_n^\epsilon(r) = (n\hat{v}_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) = (v_\epsilon/\hat{v}_\epsilon)^{1/2}[(nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_t^2)]$$

$$\to_d B(r)$$

since $\hat{v}_\epsilon \to_p v_\epsilon$ and $(nv_\epsilon)^{-1/2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) \to_d B(r)$ by Lemma 2.1(a).

(b) When $0 < \alpha < 4$ we have, as in the proof of part (b) of Theorem 3.1, the weak convergence for *fixed* r

$$a_n^{-2}\Sigma_1^{[nr]}(y_t^2 - \hat{\mu}_2) = a_n^{-2}\Sigma_1^{[nr]}(y_t^2 - \sigma_y^2) - ([nr]/n)a_n^{-2}\Sigma_1^n(y_t^2 - \sigma_y^2)$$

$$\to_d \sigma^2[U_{\alpha/2}(r) - rU_{\alpha/2}(1)] = \sigma^2K_{\alpha/2}(r) .$$

Similarly, for fixed r_1 and r_2, we have the joint weak convergence

$$\left[a_n^{-2}\Sigma_1^{[nr_1]}(y_t^2 - \hat{\mu}_2),\ a_n^{-2}\Sigma_1^{[nr_2]}(y_t^2 - \hat{\mu}_2)\right] \to_d [\sigma^2K_{\alpha/2}(r_1),\ \sigma^2K_{\alpha/2}(r_2)]$$

and the same applies to the higher finite dimensional distributions. Thus

(A30) $$a_n^{-2}\Sigma_1^{[nr]}(y_t^2 - \hat{\mu}_2) \to_{fdd} \sigma^2K_{\alpha/2}(r) .$$

Combining (A30) and (A20) we have:

$$\psi_n(r) = (n\hat{v})^{-1/2}\Sigma_1^{[nr]}(y_t^2 - \hat{\mu}_2) = \left[na_n^{-4}\hat{v}\right]^{-1/2}a_n^{-2}\Sigma_1^{[nr]}(y_t^2 - \hat{\mu}_2)$$

$$\to_{fdd} [U_{\alpha/2}(r) - rU_{\alpha/2}(1)]/\left[\int_0^1(dU_{\alpha/2})^2\right]^{1/2} ,$$

as required for (19).

To prove (20) we first write

(A31) $$\psi_n^\epsilon(r) = \left[na^{-4}\hat{v}_\epsilon\right]^{-1/2}[a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2)] .$$

By Lemma 2.1(b) we have

(A32) $$a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) \to_d K_{\alpha/2}(r) .$$

The denominator of (A31) is handled by treating the following three cases.

Case (*i*). $2 < \alpha < 4$.

(A33) $\hat{v}_\epsilon = n^{-1}\Sigma_1^n\left[\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2\right]^2 = n^{-1}\Sigma_1^n\left[\hat{\epsilon}_t^2 - \sigma_\epsilon^2\right]^2 + 2n^{-1}\Sigma_1^n(\hat{\epsilon}_t^2 - \sigma_\epsilon^2)(\sigma_t^2 - \hat{\sigma}_\epsilon^2) + \left[\sigma_\epsilon^2 - \hat{\sigma}_\epsilon^2\right]^2$

Then

$$na_n^{-4}\hat{v}_\epsilon = a_n^{-4}\Sigma_1^n\left[\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2\right]^2 + 2a_n^{-4}\Sigma_1^n(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2)(\sigma_\epsilon^2 - \hat{\sigma}_\epsilon^2) + \left[a_n^{-2}(\sigma_\epsilon^2 - \hat{\sigma}_\epsilon^2)\right]^2$$

(A34) $$\to_d \int_0^1 (dU_{\alpha/2})^2$$

since $\epsilon_t^2 \in \mathcal{ND}(\alpha/2)$ and $\hat{\sigma}_\epsilon^2 \to_p \sigma_\epsilon^2$. Using (A32) and (A34) in (A31) we obtain the required result (20).

Case (*ii*) $\alpha = 2$. In this case we recenter on $\sigma_{\epsilon n}^2 = E[\epsilon_t^2 1(\epsilon_t^2 < a_n^2)]$ in place of σ_ϵ^2 in (A33). Observe that

$$a_n^{-2}(\sigma_{\epsilon n}^2 - \hat{\sigma}_\epsilon^2) = -a_n^{-2}n^{-1}\Sigma_1^n(\hat{\epsilon}_t^2 - \sigma_{\epsilon n}^2) = O_p(n^{-1})$$

$$a_n^{-4}\Sigma_1^n\left[\hat{\epsilon}_t^2 - \sigma_{\epsilon n}^2\right]^2 = -a_n^{-4}\Sigma_1^n(\epsilon_t^2 - \sigma_{\epsilon n}^2) + o_p(1) \to_d \int_0^1 (dU_{\alpha/2})^2$$

and then

$$na_n^{-4}\hat{v}_\epsilon \to_d \int_0^1 (dU_{\alpha/2})^2$$

leading again to (20) in conjunction with (A31) and (A32).

Case (*iii*) $0 < \alpha < 2$. Here, no centering is required and we have

$$\begin{aligned} na_n^{-4}\hat{v}_\epsilon &= a_n^{-4}\Sigma_1^n\hat{\epsilon}_t^4 - na_n^{-4}\hat{\sigma}_\epsilon^4 \\ &= a_n^{-4}\Sigma_1^n\hat{\epsilon}_t^4 - n^{-1}\left[a^{-2}\Sigma_1^n\hat{\epsilon}_t^2\right] \\ &= a_n^{-4}\Sigma_1^n\hat{\epsilon}_t^4 + O_p(n^{-1}) \\ &\to_d \int_0^1 (dU_{\alpha/2})^2 \end{aligned}$$

which again leads to the required result. □

8. PROOF OF THEOREM 3.7

This is similar to Theorem 3.5. We write

$$\psi_n^\epsilon(r) = (v_\epsilon/\hat{v}_\epsilon)^{1/2} n^{-1/2}\Big\{\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) - ([nr]/n)\Sigma_1^n(\hat{\epsilon}_t^2 - \sigma_\epsilon^2)\Big\}$$

$$\text{(A35)} \qquad = (v_\epsilon/\hat{v}_\epsilon)^{1/2}\Big\{n^{-1/2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) - ([nr]/n)\Big[n^{-1/2}\Sigma_1^{n_1}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) + n^{-1/2}\Sigma_{n_1+1}^n(\hat{\epsilon}_t^2 - \sigma_\epsilon^2)\Big] + n_2 n^{-1/2}(\sigma_{\epsilon+}^2 - \sigma_\epsilon^2)\Big\} .$$

Now suppose that $\alpha > 4$. Since $v_\epsilon/\hat{v}_\epsilon = O_p(1)$ and $\sigma_\epsilon^2 \neq \sigma_{\epsilon+}^2$ under H_1 it is easy to see that all the finite dimensional distributions of $\psi_n^\epsilon(r)$ diverge under H_1. For $r \le r_k = k/(1+k)$ we have from (A35) that $\psi_n^\epsilon(r) = O_p(n_2 n^{-1/2}) = O_p(n^{1/2})$ as $n \to \infty$. A similar decomposition for $\psi_n^\epsilon(r)$ applies when $r > r_k$ leading to the same rate of divergence over the interval $(r_k, 1]$ and we have (21) as required.

When $2 < \alpha < 4$ we have in place of (A35) the decomposition

$$a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) = a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) + nr]a_n^{-2}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2)$$

$$= a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) - ([nr]/n)a_n^{-2}\Big[\Sigma_1^{n_1}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) + \Sigma_{n_1+1}^{n_2}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2) + n_2(\sigma_{\epsilon+}^2 - \sigma_\epsilon^2)\Big] .$$

Then, taking $r \le r_k$ we have

$$\psi_n^\epsilon(r) = (na_n^{-4}\hat{v}_\epsilon)^{-1/2} a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \sigma_\epsilon^2)$$

$$= O_p(n_2 a_n^{-2}) = O_p(n^{1-2/\alpha})$$

as required for (22). Again the same rate of divergence applies when $r > r_k$.

Finally, when $\alpha < 2$ we write

$$a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) = a_n^{-2}\Sigma_1^{[nr]}\hat{\epsilon}_t^2 - ([nr]/n)a_n^{-2}\Sigma_1^n\hat{\epsilon}_t^2$$

$$= a_n^{-2}\Sigma_1^{[nr]}\hat{\epsilon}_t^2 - ([nr]/n)\Big\{a_n^{-2}\Sigma_1^{n_1}\hat{\epsilon}^2 + a_n^{-2}\Sigma_{n_1+1}^n\hat{\epsilon}_t^2\Big\}$$

$$= O_p(1) ,$$

and $na_n^{-4}\hat{v}_\epsilon = O_p(1)$ as in the proof of Theorem 3.5. Thus, we find

$$\psi_n^\epsilon(r) = (na_n^{-4}\hat{v}_\epsilon)^{-1/2} a_n^{-2}\Sigma_1^{[nr]}(\hat{\epsilon}_t^2 - \hat{\sigma}_\epsilon^2) = O_p(1)$$

and the test is inconsistent in this case.

9. PROOF OF THEOREM 3.8

Using (26) we have

$$s^{1/2}[\ln(n_1/s)]^{-1}\tau_d = s^{1/2}[\ln(n_1/s)]^{-1}(\hat{d}_s^{(1)} - d) - s^{-1/2}\ln(n_1/s)]^{-1}(\hat{d}_s^{(2)} - d)$$
$$\to_d N(0, 2d^2)$$

and $$\tilde{V}(\tau_d) = s^{1/2}[\ln(n_1/s)]^{-1}\tau_d/\left[\hat{d}_s^{(1)^2} - \hat{d}_s^{(2)^2}\right]^{1/2} \to_d N(0,1)$$

as required. Note that we may proceed as if $\hat{\epsilon}_{n,j}$ is replaced by $\epsilon_{n,j}$ in (23) and (24) since $\hat{\epsilon}_t$ (and hence $\hat{\epsilon}_{n,j}$) is consistent for ϵ_t (respectively $\epsilon_{n,j}$) under both the null and alternative hypotheses.

10. PROOF OF THEOREM 3.9

Under H_1 we have

$$s^{1/2}[\ln(n_1/s)]^{-1}\tau_d = s^{1/2}[\ln(n_1/s)]^{-1}(\hat{d}_s^{(1)} - d^{(1)}) - s^{1/2}[\ln(n_1/s)]^{-1}(\hat{d}_s^{(2)} - d^{(2)})$$
$$+ s^{1/2}[\ln(n_1/s)]^{-1}(d^{(1)} - d^{(2)})$$
$$= O_p(s^{1/2}[\ln(n_1/s)]^{-1}).$$

The same divergence rate applies to $\tilde{V}(\tau_d)$ as $n \to \infty$ since $\hat{d}_s^{(1)} \to_p d^{(1)}$ and $\hat{d}_s^{(2)} \to_p d^{(2)}$.

11. PROOF OF THEOREM 3.10

Using (25) we have under H_0^α

$$s^{1/2}(\hat{\alpha}_s^{(1)} - \hat{\alpha}_s^{(2)}) = s^{1/2}(\hat{\alpha}_s^{(1)} - \alpha) - s^{1/2}(\hat{\alpha}_s^{(2)} - \alpha) \to_d N(0, 2\alpha^2)$$

and (30) follows directly. Similarly under H_1^α we have

$$s^{1/2}(\hat{\alpha}_s^{(1)} - \hat{\alpha}_s^{(2)}) = s^{1/2}(\hat{\alpha}_s^{(1)} - \alpha^{(1)}) - s^{1/2}(\hat{\alpha}_s^{(2)} - \alpha^{(2)}) + s^{1/2}(\alpha^{(1)} - \alpha^{(2)}) = O_p(s^{1/2})$$

as required for (31).

DATA APPENDIX

In this section we describe the transformations that were applied to the data prior to testing for covariance stationarity under moment condition failure. The purpose of these preliminary transformations is not estimation of structural parameters of the DGP's, but merely calculation of the unconditional variances while taking into account regularities in the means of the series. We also need to eliminate serial dependence in the mean in order to apply the asymptotic distribution theory developed for the cusum of squares test when fourth moments are infinite. For simplicity, we do not explicitly consider long memory processes, and in eliminating time series effects the maintained hypothesis will be that the data are covariance stationary. This hypothesis is tested at length, of course, in section 5 of our paper.

The first series we study are monthly returns to an aggregate stock market portfolio which covers the period from January 1834 to December 1987 $(n = 1{,}848)$. Schwert (1989) discusses the details of the construction of this series. The data were very generously provided to us by Adrian Pagan and Bill Schwert, and were furnished as the residuals from a regression of the returns series on 12 monthly dummies. This transformation demeans the data and removes seasonal effects at monthly frequencies. An analysis of the serial dependence in this series led to specification and estimation of the following AR(12) process:

$$\hat{y}_t = \Sigma_{i=1}^{12} \hat{b}_i y_{t-i} \quad (t = 1835{:}1, \ldots, 1987{:}12).$$

The estimated coefficients and their standard errors are:

i	$\hat{b}_i$	$\hat{s}_i$
1	0.1630	.00104
2	−0.0219	.0234
3	−0.0687	.0238
4	0.0305	.0239
5	0.0655	.0239
6	−0.0317	.0240
7	0.0142	.0240
8	0.0427	.0239
9	0.0413	.0239
10	0.0183	.0239
11	0.0099	.0239
12	0.0037	.0236

$$R^2 = 0.0414\ ,\ \bar{R}^2 = 0.0351\ ,\ F_{12,1824} = 6.56\ .$$

The largest estimated coefficient is that of the first order AR component, $\hat{b}_1 = 0.163$, but far more than just one component had to be included in an AR model in order to remove serial dependence from the data. The Q–Statistic at lag length 20 is equal to 31.7 and serves to test whether the first 20 autocorrelation coefficients of the residuals from the AR(12) model are jointly zero; the statistic is distributed asymptotically as a χ^2 variate with 20 degrees of freedom under the null of no serial correlation. Since the test statistic is below the corresponding 95% critical value, no further transformations were applied to the data, and this residual series is the series of transformed monthly stock returns which we use in section 5 for the analysis of covariance stationarity of monthly stock returns. The correlation coefficient between the residual series and the original series is 0.979. Since large fluctuations tend to be serially correlated empirically, a consequence of regressing out an AR(12) process from the data is to reduce the largest fluctuations slightly in magnitude. For instance, the largest positive observation was reduced from 36.8% (the unit of measurement is percentage change per month) to 33.8%, and the largest negative deviation was reduced from −28.6% to −26.6%. The largest fluctuations all occurred during the Great Depression.

The second series is the sequence of daily returns to the Standard and Poors 500 stock market index. The series is measured in percentage changes per day and covers the time span from 7/2/1962 to 12/31/1987 ($n = 6{,}409$). The data were obtained from the 1988 CRSP tape. A regression of the returns series on 5 weekday dummies gave the following estimates:

i	$\hat{b}_i$	$\hat{s}_i$
1	−0.00135	.00025
2	0.00034	.00024
3	0.00110	.00024
4	0.00040	.00024
5	0.00080	.00025

$$R^2 = 0.00911\ ,\ \bar{R}^2 = 0.00849\ ,\ F_{5,6404} = 14.7\ .$$

The mean of the dependent variable is equal to 0.00027, thus all coefficient estimates are of the same order of magnitude as the mean of the series itself. Mondays (or, rather, holding the stock index over weekends, from Friday evening to Monday evening) appear to generate negative returns, while the other weekdays show positive expected returns. Overall, though, the day–of–week effects contribute little to explaining (statistically) the fluctuations of the daily data (cf. the low value of the R^2 statistic). Next, in order to remove additional calendar effects from the data, we regressed the residuals from the first regression on 12 monthly dummy variables. The estimates from the second regression are not reported here; all estimated coefficients were insignificant according to conventional t–tests and the regression F–test and were much smaller in magnitude than the coefficients of the weekday dummies shown above. Finally, we analyzed the time series properties of the residuals from the two consecutive calendar effect regressions. We specified an AR(5) process for the residuals, the coefficient estimates and corresponding standard errors for this choice of model are:

i	$\hat{b}_i$	$\hat{s}_i$
1	0.1602	.0125
2	−0.0502	.0127
3	0.0118	.0127
4	−0.0264	.0127
5	0.0267	.0125

$$R^2 = 0.0265\ ,\ \bar{R}^2 = 0.0259\ ,\ F_{5,6404} = 43.6\ .$$

As is the case for the monthly data, the first order autoregressive coefficient is the largest. Including 5 lags in the AR model proved to be sufficient to eliminate significant time series effects from the data; the Q–statistic for zero serial correlation up to 20 lags is equal to 26.8 for the residuals from this AR(5) model. The null hypothesis of zero serial correlation could thus not be rejected, and this residual series is used in section 5 to test for covariance stationarity of the daily stock returns. The correlation coefficient between the original Standard and Poors 500 series and the residuals series is equal to 0.969. The elimination of the time series effects again led to a slight smoothing of the large fluctuations; e.g., the largest negative return (which was observed on October 19, 1987) was reduced from –20.5% to –19.5%.

Table 1. Critical Values of Sample Split Prediction Test Statistic $V_1(\tau)$

P(X < c)	2.1	2.5	α 3.0	3.5	3.8	N(0,1) [α>4]
90%	1.26	1.28	1.28	1.29	1.28	1.282
95%	1.51	1.55	1.59	1.63	1.62	1.645
97.5%	1.73	1.79	1.85	1.91	1.93	1.960
99%	1.99	2.07	2.15	2.24	2.28	2.326
99.5%	2.17	2.26	2.34	2.48	2.51	2.576

For $\alpha < 4$, critical values are based on 50,000 simulations of the test statistic, with a sample size of $n = 1{,}000$. For $\alpha > 4$, standard normal critical values apply. Note: In Tables 1–4, all critical values are for one–sided tests of the corresponding null hypotheses.

Table 2. Critical Values of Finite Dimensional Distributions (fdd) of the Cusum of Squares Test Statistics $\psi_n(r)$ and $\psi_n^\epsilon(r)$

(2.a) $\alpha = 2.1$

P(X < c)	r 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
50%	−0.10	−0.13	−0.11	−0.06	0.00	0.06	0.11	0.13	0.10
90%	0.39	0.61	0.66	0.63	0.59	0.53	0.46	0.36	0.23
95%	0.66	0.78	0.75	0.74	0.70	0.63	0.54	0.43	0.27
97.5%	0.84	0.85	0.86	0.84	0.78	0.71	0.61	0.48	0.30
99%	0.91	0.98	0.98	0.95	0.89	0.80	0.69	0.54	0.34
99.5%	0.98	1.07	1.07	1.02	0.96	0.87	0.74	0.58	0.36

(2.b) $\alpha = 2.5$

P(X < c)	r 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
50%	−0.09	−0.11	−0.09	−0.05	0.00	0.05	0.09	0.11	0.09
90%	0.43	0.62	0.68	0.66	0.64	0.59	0.52	0.42	0.28
95%	0.67	0.79	0.81	0.80	0.77	0.72	0.63	0.51	0.34
97.5%	0.83	0.90	0.94	0.93	0.89	0.83	0.73	0.59	0.40
99%	0.93	1.05	1.08	1.07	1.03	0.96	0.84	0.69	0.46
99.5%	1.02	1.14	1.19	1.18	1.13	1.05	0.93	0.75	0.51

(2.c) $\alpha = 3.0$

P(X < c)	r 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
50%	−0.06	−0.07	−0.06	−0.03	0.00	0.03	0.06	0.07	0.06
90%	0.41	0.58	0.65	0.65	0.64	0.61	0.54	0.45	0.32
95%	0.61	0.76	0.80	0.81	0.79	0.75	0.67	0.55	0.39
97.5%	0.78	0.88	0.94	0.95	0.93	0.88	0.78	0.65	0.45
99%	0.90	1.02	1.10	1.11	1.08	1.03	0.91	0.76	0.53
99.5%	0.99	1.12	1.21	1.22	1.19	1.12	1.00	0.83	0.58

(Table 2, continued)

(2.d) $\alpha = 3.5$

P(X < c)	r = 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
50%	−0.04	−0.04	−0.03	−0.02	0.00	0.02	0.03	0.04	0.04
90%	0.40	0.55	0.62	0.64	0.64	0.61	0.57	0.48	0.35
95%	0.55	0.72	0.79	0.81	0.81	0.77	0.71	0.61	0.44
97.5%	0.69	0.85	0.93	0.95	0.95	0.91	0.84	0.71	0.52
99%	0.85	1.00	1.10	1.13	1.12	1.07	0.99	0.84	0.61
99.5%	0.93	1.10	1.20	1.25	1.23	1.18	1.09	0.92	0.67

(2.e) $\alpha = 3.8$

P(X < c)	r = 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
50%	−0.01	−0.01	−0.01	−0.01	0.00	0.01	0.01	0.01	0.01
90%	0.39	0.52	0.60	0.64	0.64	0.63	0.58	0.50	0.37
95%	0.52	0.68	0.77	0.81	0.82	0.80	0.74	0.64	0.47
97.5%	0.63	0.81	0.91	0.96	0.97	0.95	0.88	0.76	0.56
99%	0.77	0.96	1.08	1.13	1.15	1.12	1.03	0.89	0.66
99.5%	0.86	1.06	1.20	1.24	1.27	1.24	1.15	0.99	0.73

(2.f) $\alpha > 4.0$ (exact critical values)

P(X < c)	r = 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
50%	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
90%	0.38	0.51	0.59	0.63	0.64	0.63	0.59	0.51	0.38
95%	0.49	0.66	0.75	0.81	0.82	0.81	0.75	0.66	0.49
97.5%	0.59	0.78	0.90	0.96	0.98	0.96	0.90	0.78	0.59
99%	0.70	0.93	1.07	1.14	1.16	1.14	1.07	0.93	0.70
99.5%	0.77	1.03	1.18	1.26	1.29	1.26	1.18	1.03	0.77

The critical values in Tables 2.a–2.e are based on 50,000 simulations of the test statistics, with a sample size of $n = 1{,}000$ (except Table 2.a: $n = 2{,}500$). The critical values in Table 2.f are the exact critical values of the Brownian bridge statistic, calculated from $B(r) =_d N(0, r(1-r))$.

Table 3. Critical Values of the $\sup_r(\psi_n^{\epsilon}(r)$ Statistic

P(X < c)	2.1	2.5	α 3.0	3.5	3.8	N(0,1) [α>4]
80%	0.76	0.83	0.85	0.86	0.87	0.897
90%	0.89	0.97	1.00	1.02	1.04	1.073
95%	0.98	1.09	1.13	1.17	1.19	1.224
97.5%	1.10	1.24	1.29	1.33	1.36	1.358
99%	1.18	1.34	1.40	1.44	1.48	1.517

The critical values are based on 50,000 simulations of the test statistics for $\alpha < 4$, with a sample size of 1,000 (except for $\alpha = 2.1$: $n = 2{,}500$). For $\alpha > 4$, the exact critical values c solve the equation $P(\sup_r(B(r))>c) = \exp(-2c^2)$, $c > 0$ (Billingsley (1968), equation (11.40), p. 85).

Table 4. Critical Values of the R_L Statistic

P(X < c)	2.1	2.5	α 3.0	3.5	3.8	N(0,1) [α>4]
80%	1.13	1.27	1.32	1.37	1.41	1.473
90%	1.23	1.39	1.45	1.51	1.55	1.620
95%	1.31	1.50	1.57	1.63	1.68	1.747
97.5%	1.41	1.63	1.71	1.77	1.83	1.862
99%	1.48	1.72	1.80	1.87	1.93	1.961

The critical values are based on 50,000 simulations of the test statistics for $\alpha < 4$, with a sample size of $n = 1000$ (except in the case of $\alpha = 2.1$: $n = 2{,}500$). For $\alpha > 4$, exact critical values can be computed; the figures reproduced here are from Haubrich and Lo (1989, Table 1a).

Table 5. Point estimates and standard errors of the maximal moment exponent α of the empirical distributions

(5.a) Monthly stock returns

s	left tail (n = 922) $\hat{\alpha}$	(s.e.)	right tail (n = 914) $\hat{\alpha}$	(s.e.)
20	3.55	(0.79)	2.95	(0.66)
40	3.12	(0.49)	2.46	(0.39)
60	3.22	(0.42)	2.45	(0.32)
80	3.00	(0.34)	2.61	(0.29)
100	2.95	(0.29)	2.66	(0.27)

(5.b) Daily stock returns

s	left tail (n = 3,197) $\hat{\alpha}$	(s.e.)	right tail (n = 3,207) $\hat{\alpha}$	(s.e.)
50	3.80	(0.54)	3.37	(0.48)
100	3.79	(0.38)	3.86	(0.39)
150	3.59	(0.29)	3.44	(0.28)
200	3.68	(0.26)	3.17	(0.22)
250	3.44	(0.22)	3.08	(0.19)

Remark: See text for the formulas to compute point estimates and standard errors of α .

Table 6. **Tests of constancy of tail parameter α across tails and over time ($\tilde{V}(\tau_\alpha)$ test)**

(6.a) Monthly stock returns

s	Constancy of α across tails	Constancy of α over time
40	1.045	0.865
60	1.467	0.886
80	0.882	1.348

(6.b) Daily stock returns

s	Constancy of α across tails	Constancy of α over time
100	−0.124	0.055
150	0.369	−0.477
180	1.385	−0.398

Remark: Use standard normal critical values for this test statistic. The sample split prediction tests reported in the second column were performed on the absolute values of the returns, in order to work with approximately the same number of observations as in the first column, where the split was between positive and negative values.

Figure 1
Recursive Variance

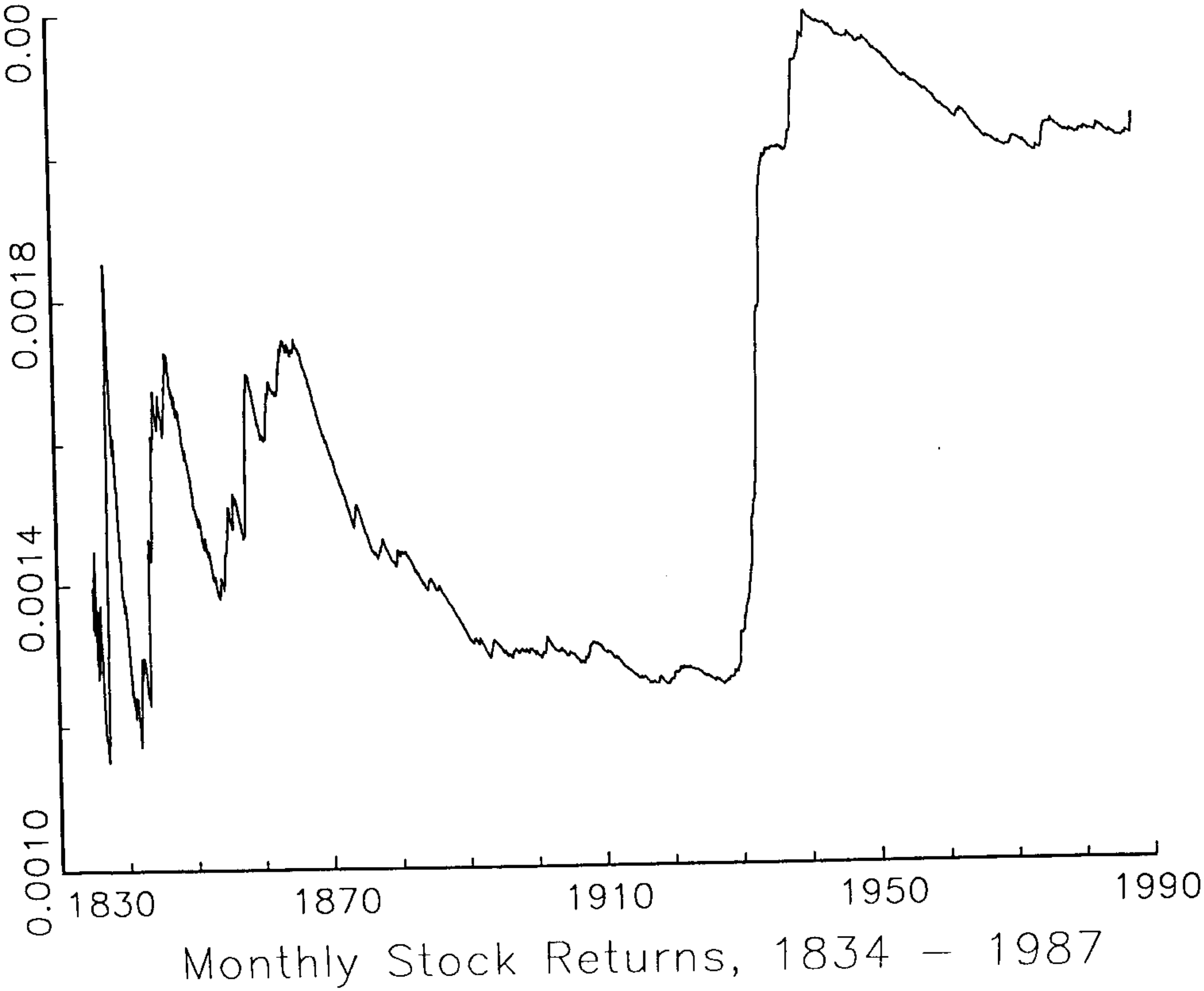
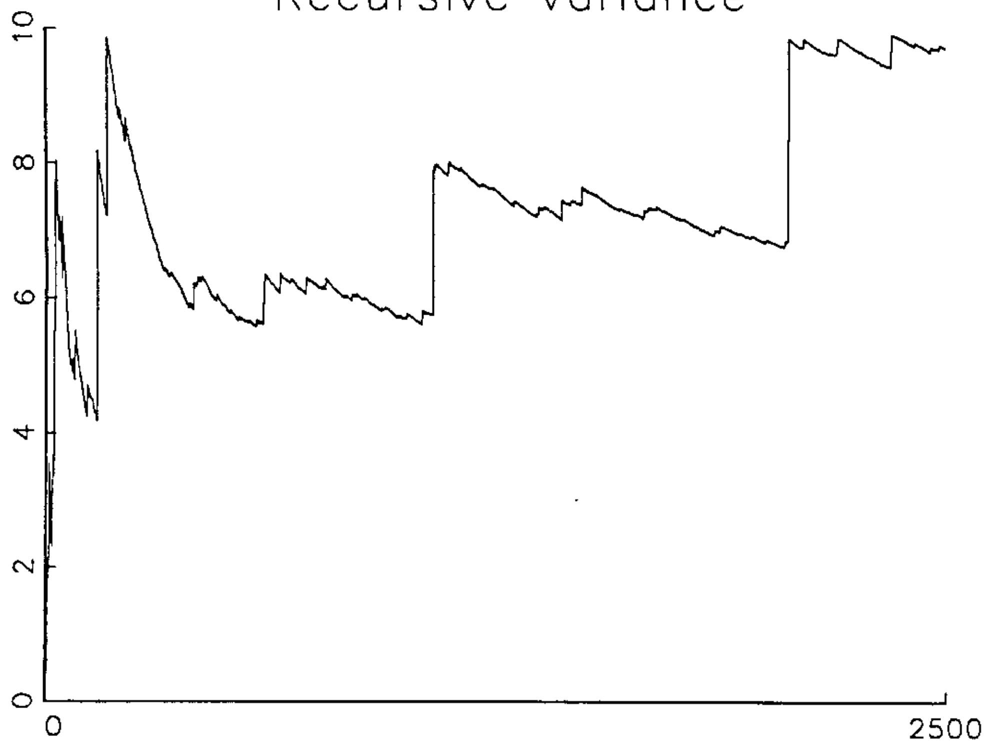
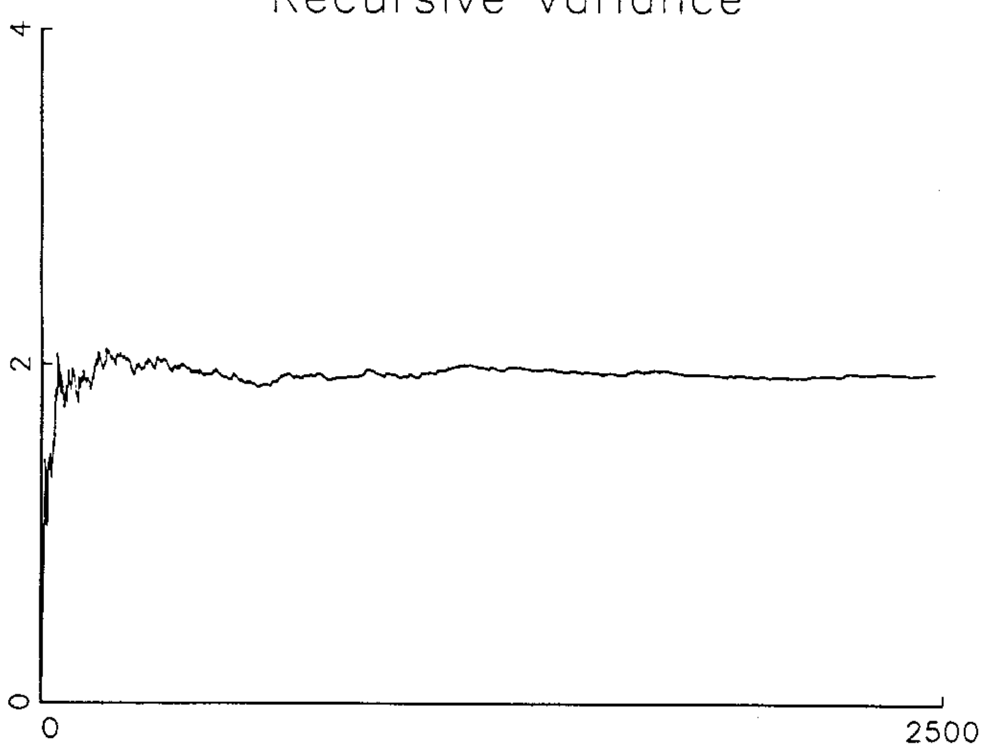


Figure 2
Recursive Variance



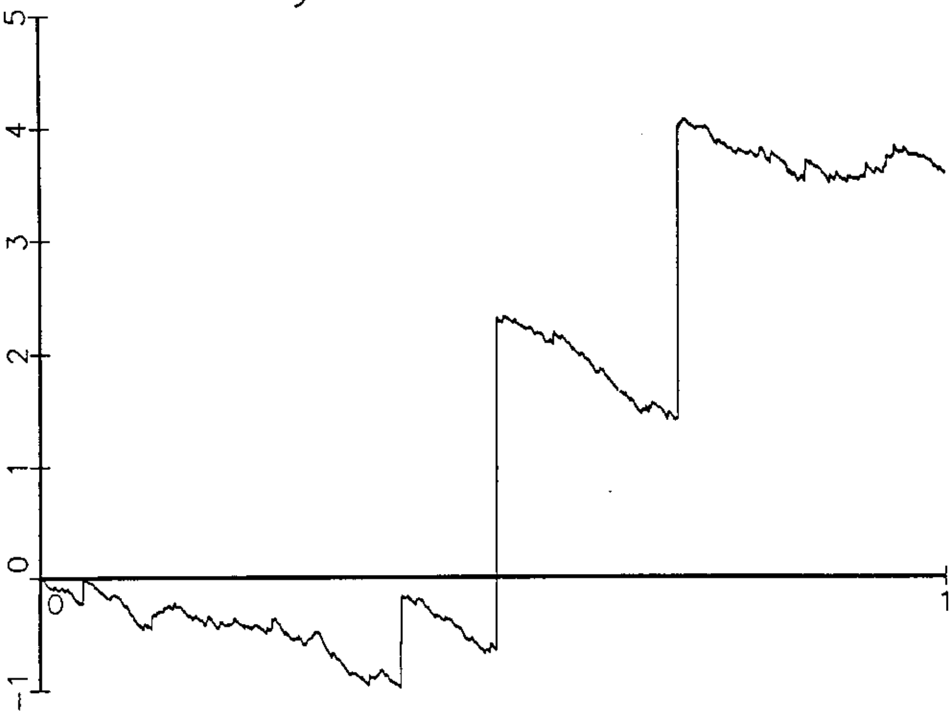
iid symm. stable variates, $\alpha=1.5$, n=2500

Figure 3
Recursive Variance



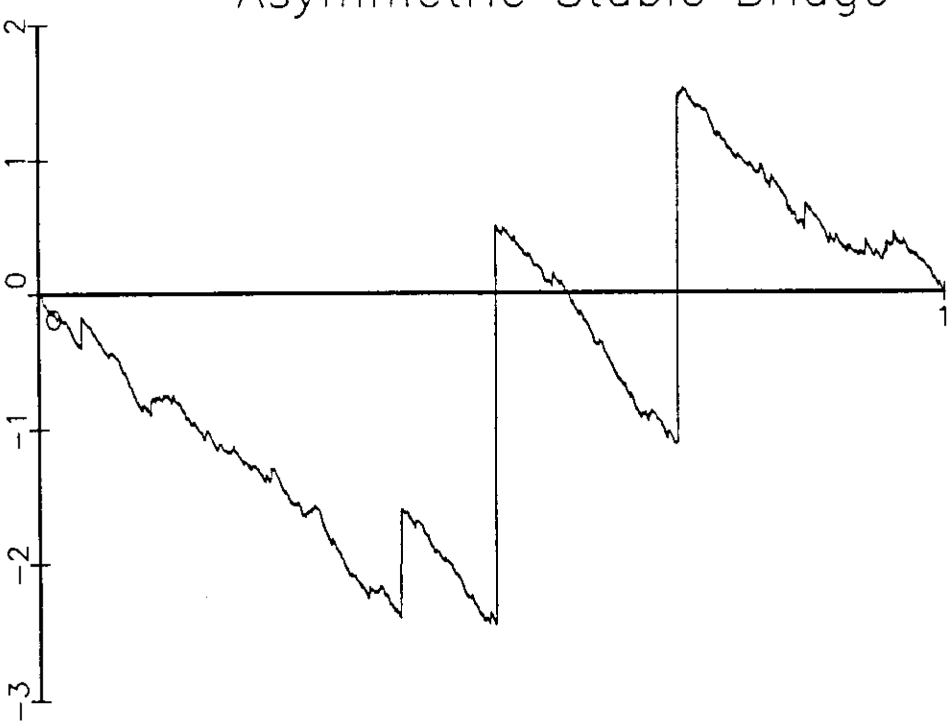
iid Gaussian variates, $\sigma^2=2$, n=2500

Figure 4a
Asymmetric Stable Process



n = 2500, $\alpha/2$ = 1.5

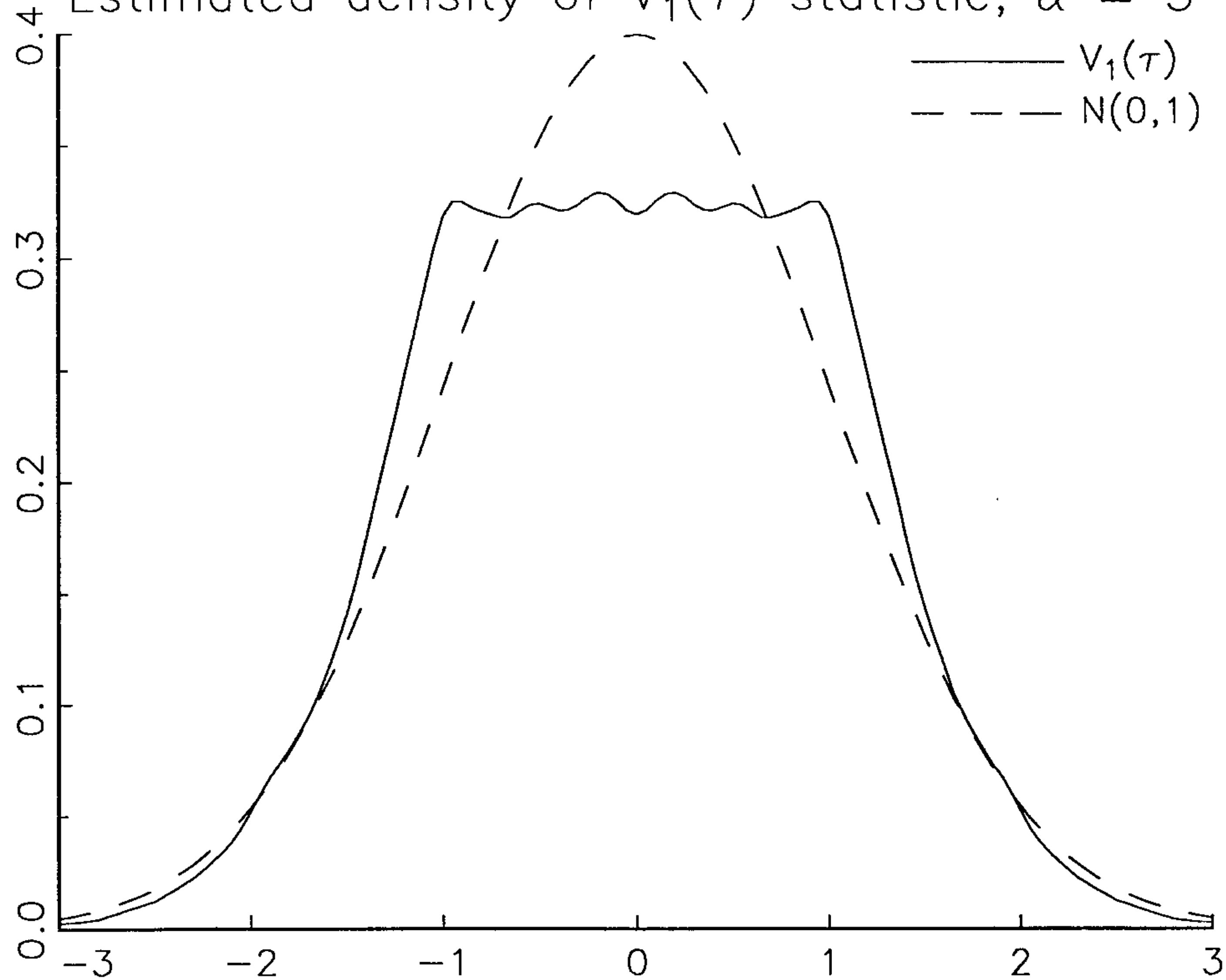
Figure 4b
Asymmetric Stable Bridge



n = 2500, $\alpha/2$ = 1.5

Figure 5

Estimated density of $V_1(\tau)$ statistic, $\alpha = 3$



sample size: 2,500; 50,000 iterations; bandwidth parameter h=0.06

Figure 6a
Cusum^2 test: fdd bounds, α = 2.1

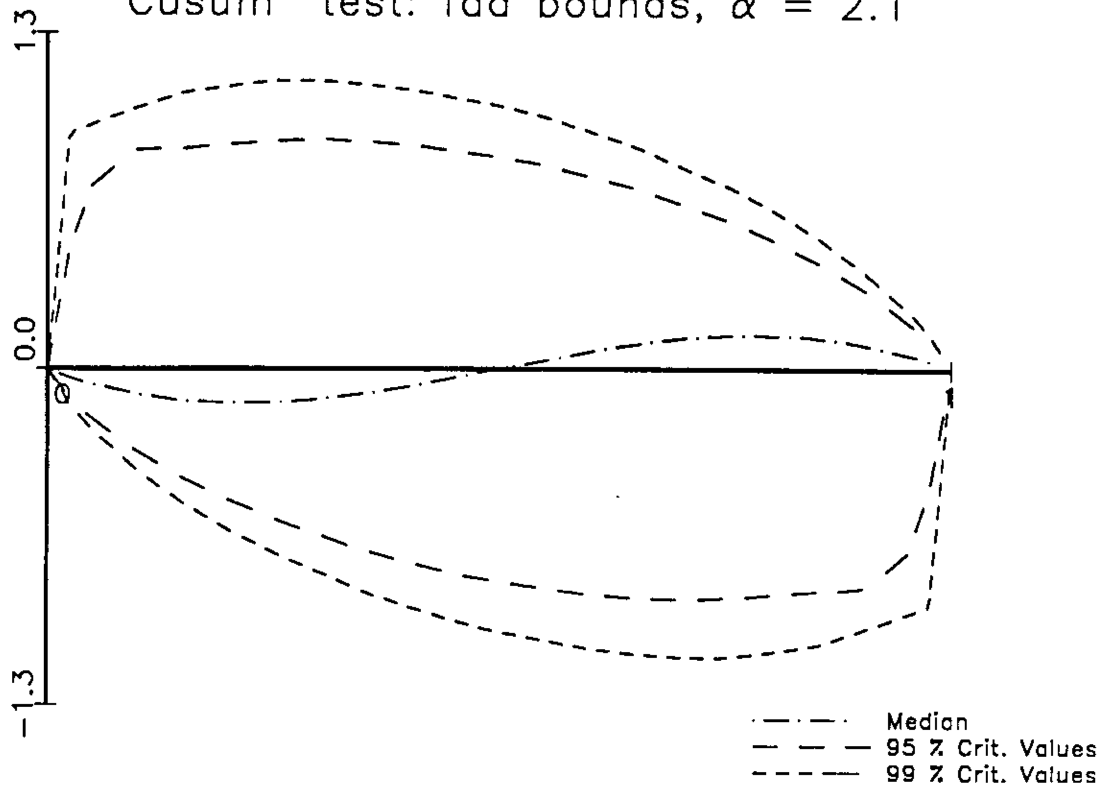


Figure 6b
Cusum^2 test: fdd bounds, α = 2.5

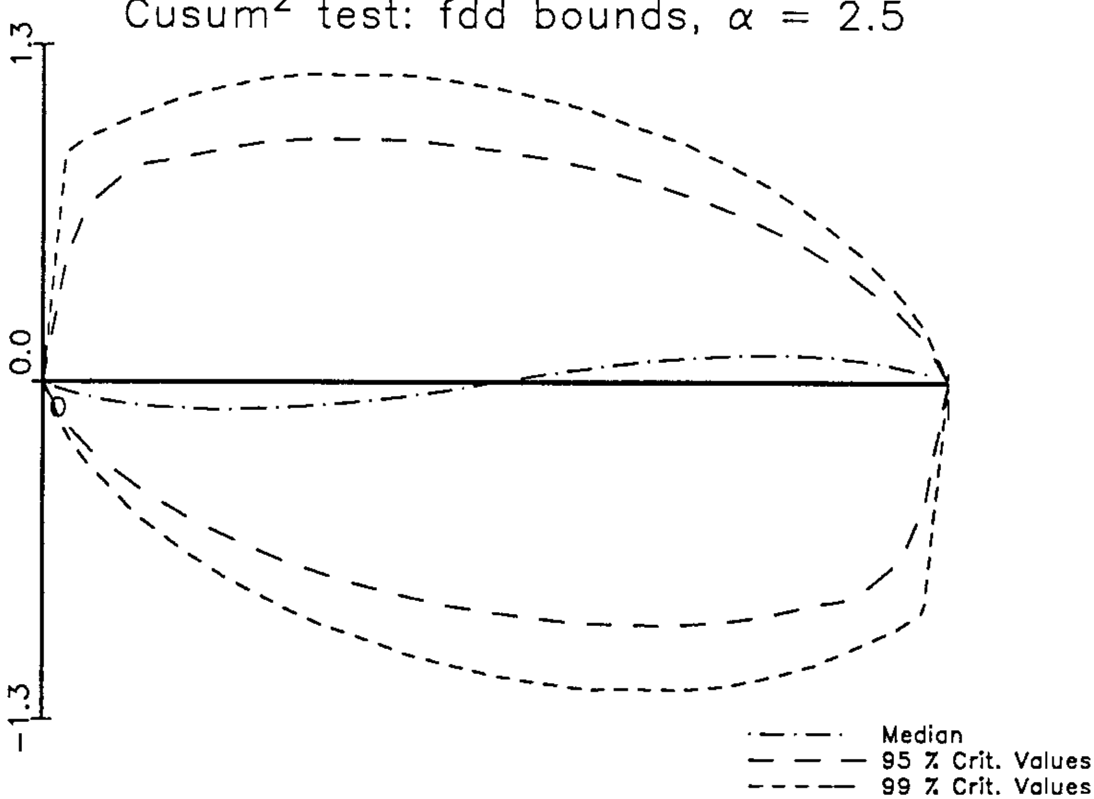


Figure 6c

Cusum^2 test: fdd bounds, α = 3.0

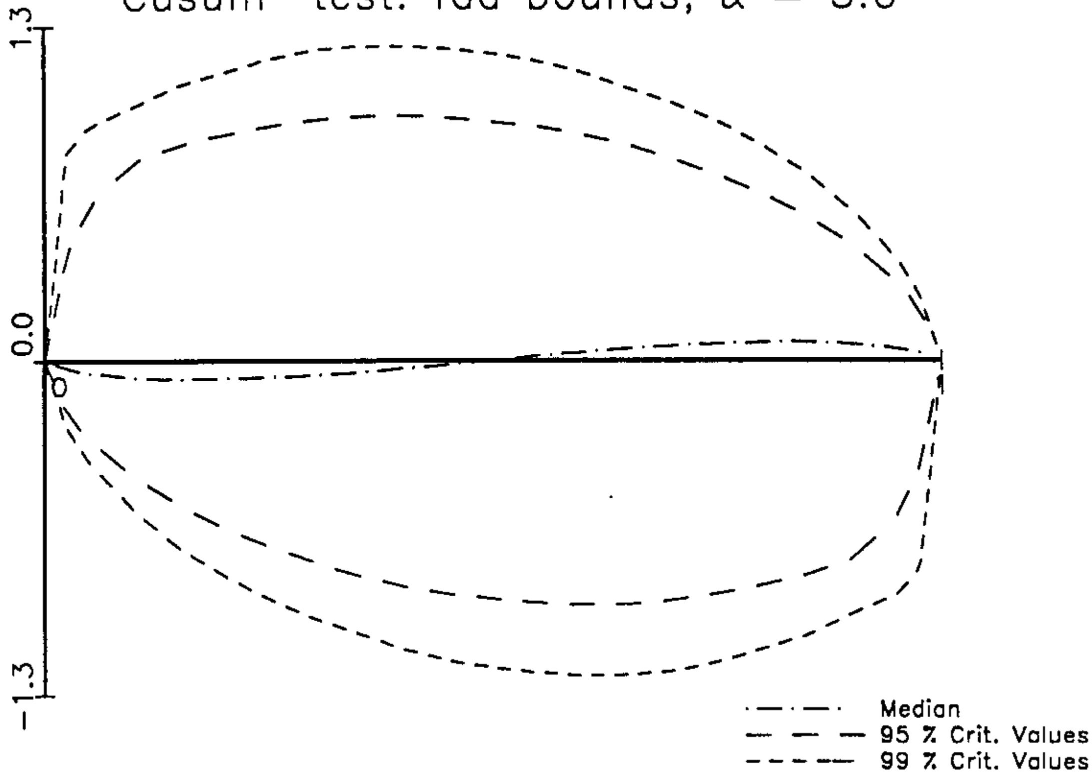


Figure 6d

Cusum^2 test: fdd bounds, α = 3.5

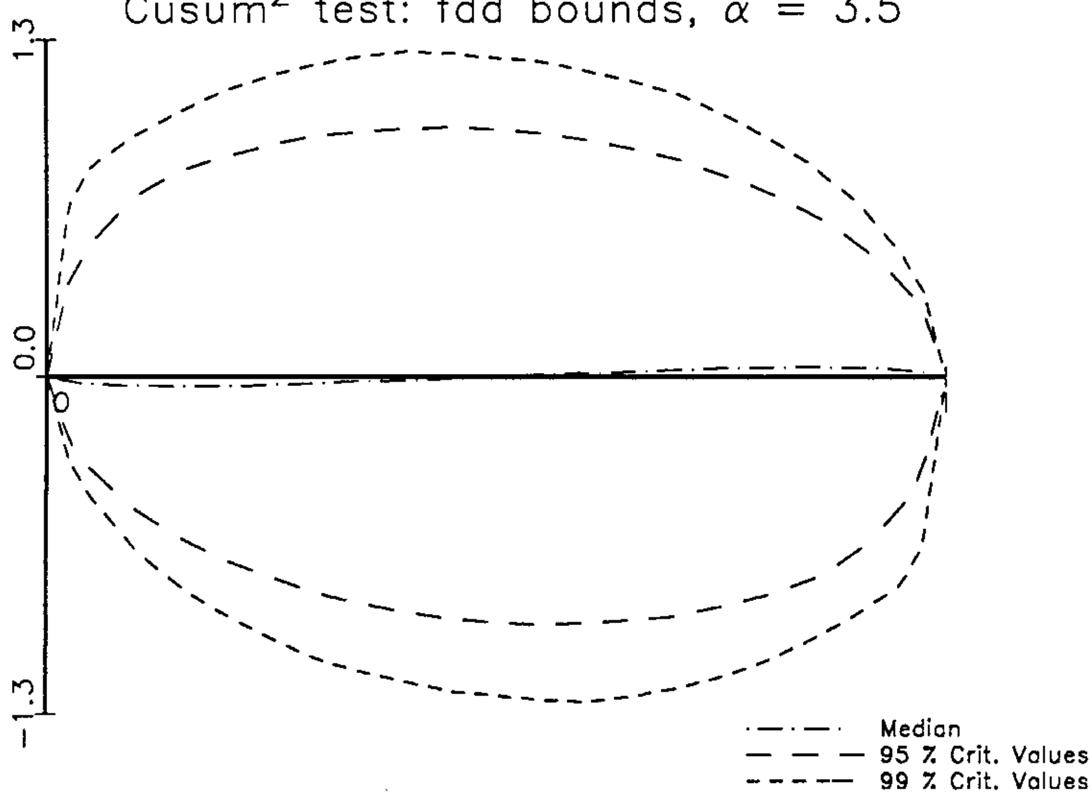


Figure 6e

Cusum^2 test: fdd bounds, α = 3.8

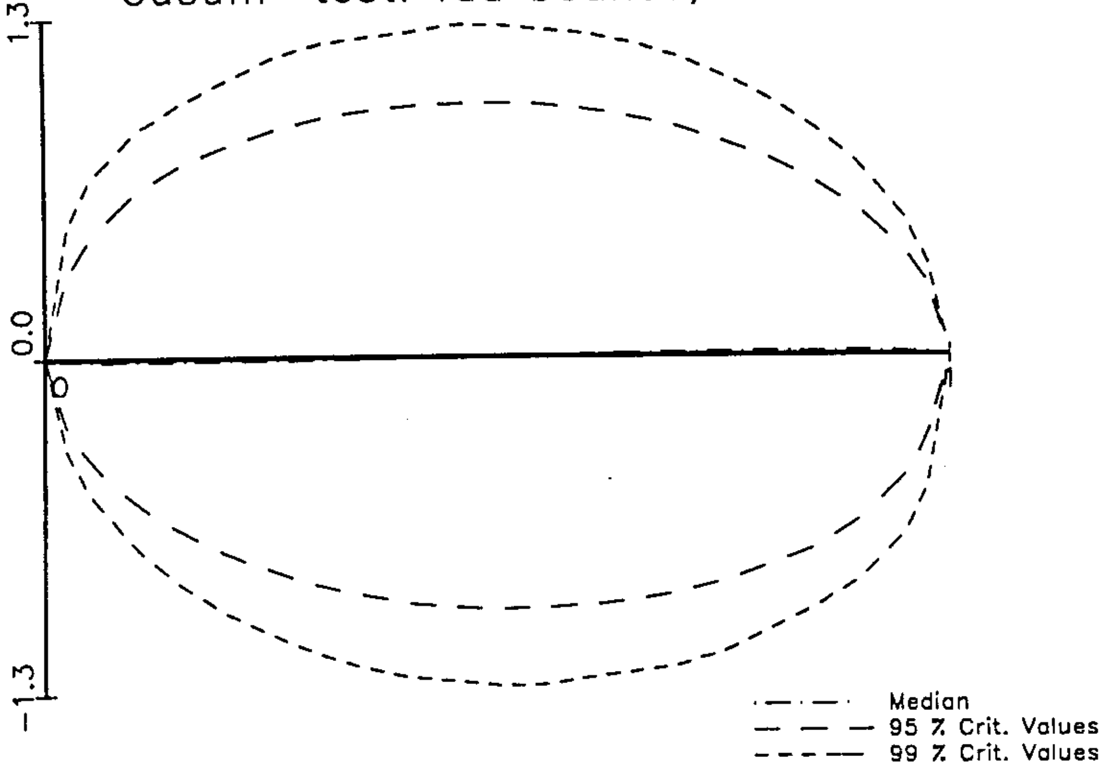


Figure 6f

Cusum^2 test: fdd bounds, α > 4.0

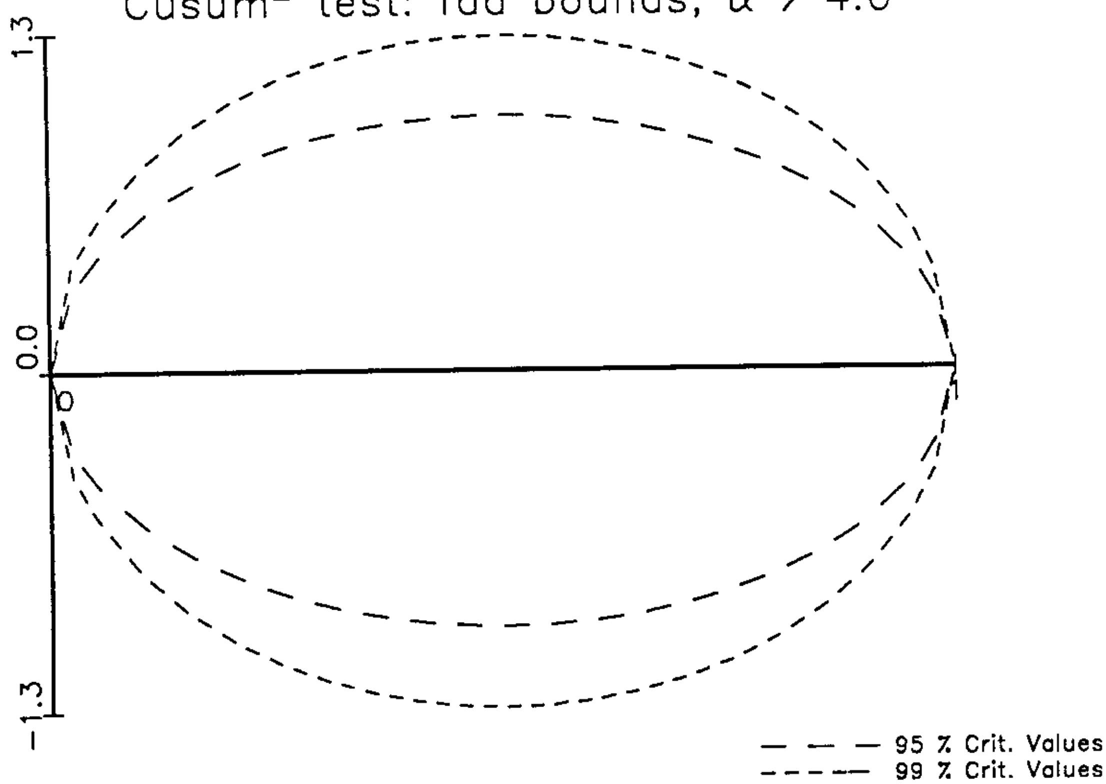


Figure 7a

99 % conf. contours, cusum^2 test

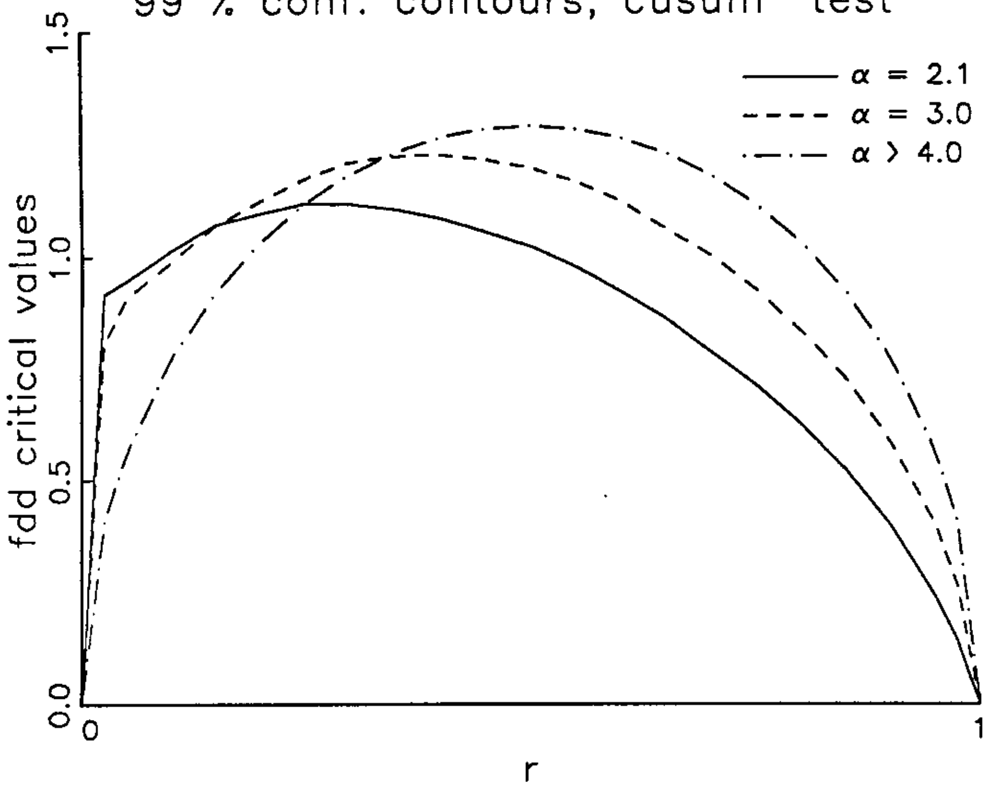


Figure 7b

95 % conf. contours, cusum^2 test

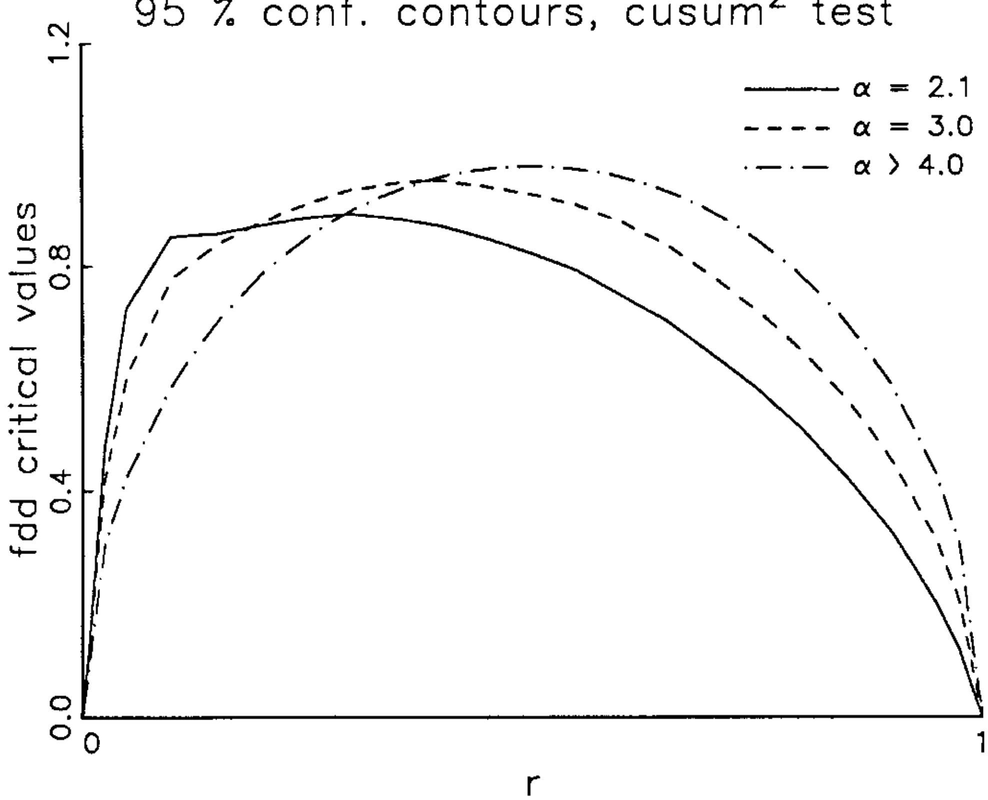
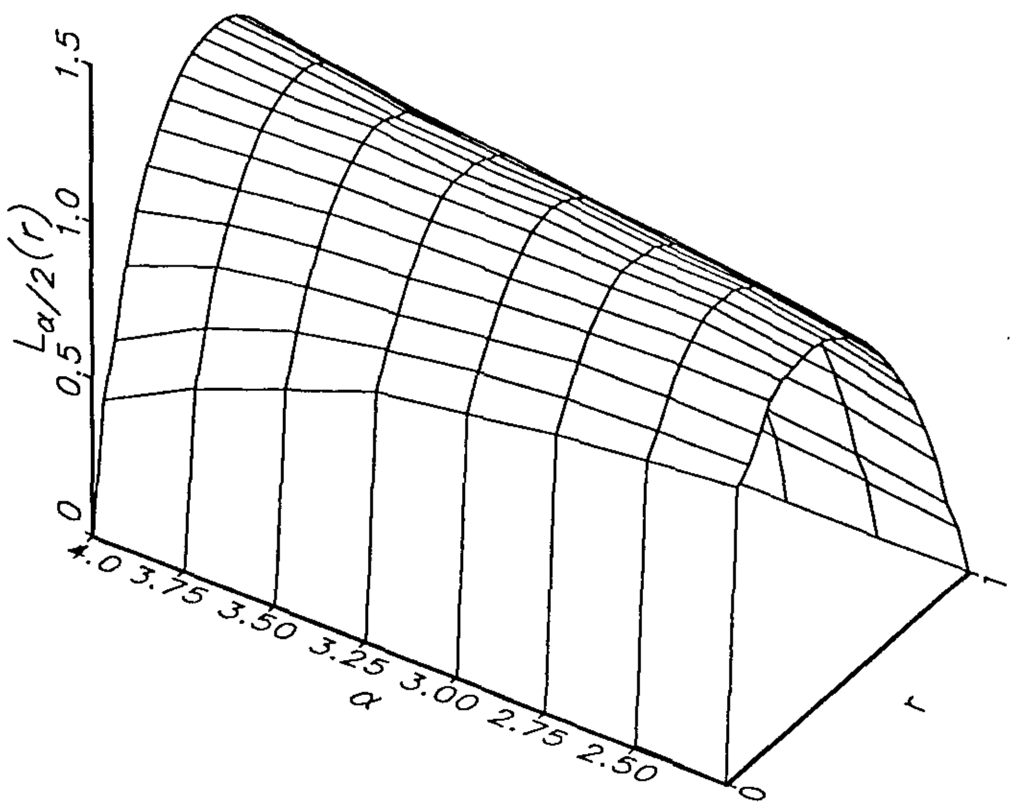


Figure 8

Cusum^2 test: 99% fdd bounds



$2.25 \leqq \alpha \leqq 4.0, \; 0 \leqq r \leqq 1$

Figure 9a

cdf, Monthly Stock Returns

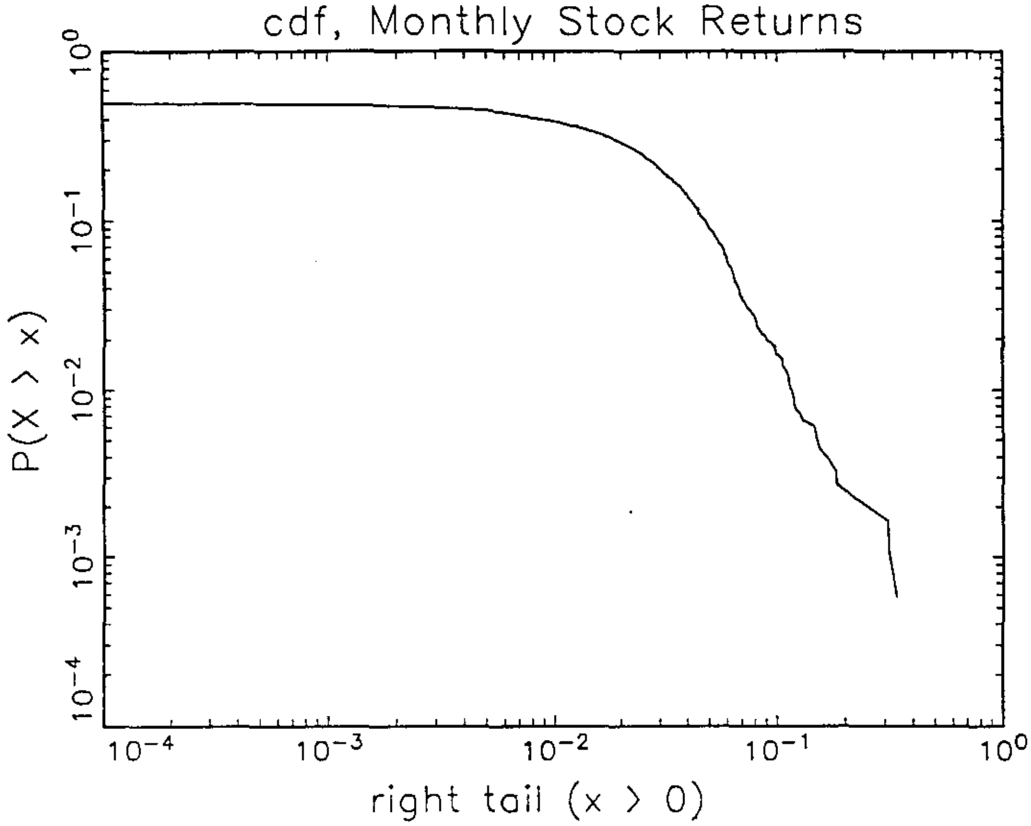


Figure 9b

cdf, Daily Stock Returns

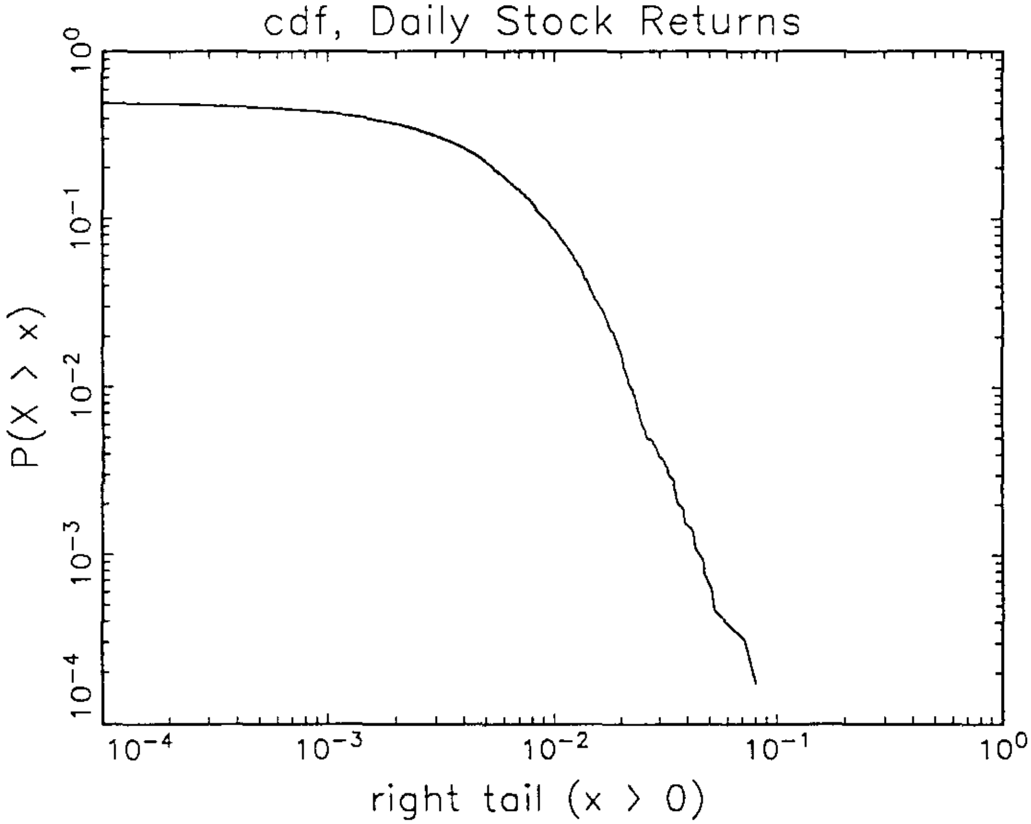


Figure 10a

Cusum² test, monthly stock returns series

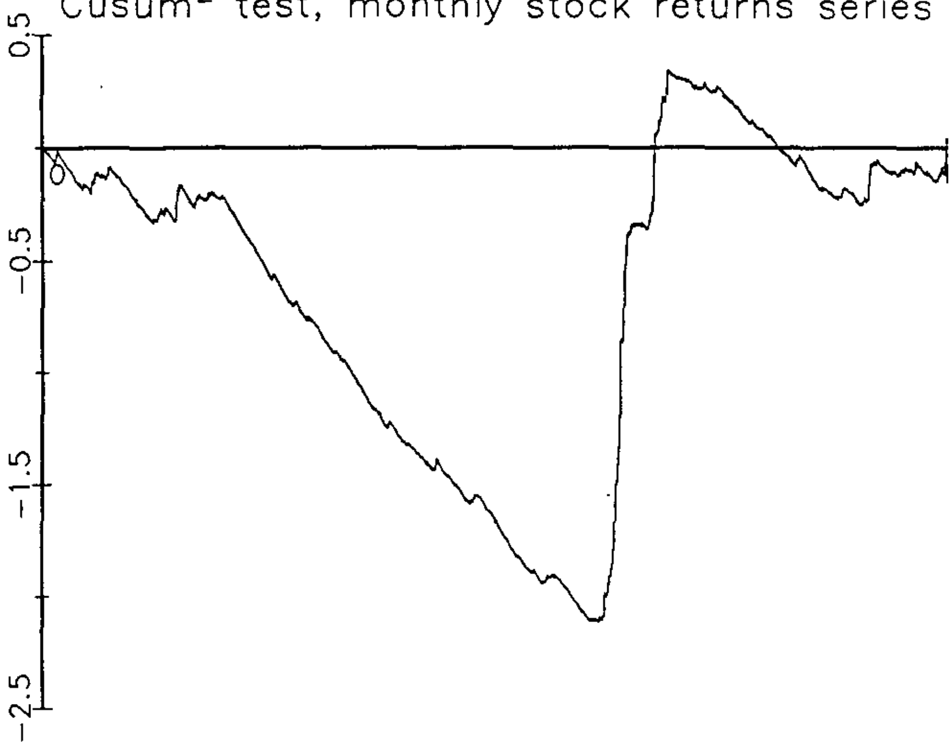
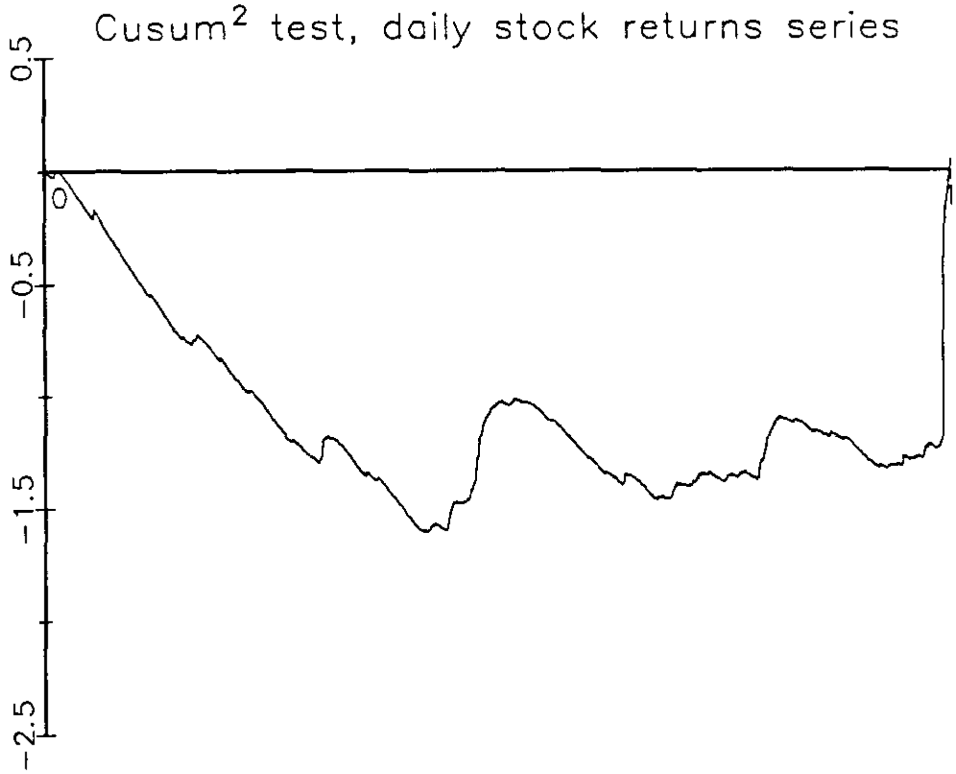


Figure 10b

Cusum² test, daily stock returns series



REFERENCES

Akgiray, V. and G.G. Booth (1988), "The stable–law model of stock returns," *Journal of Business and Economic Statistics* 6, 51–57.

Andrews, D.W.K. (1986), "On the performance of least squares in linear regression with undefined error means," Cowles Foundation Discussion Paper No. 798 (May).

Avram, F. and M.S. Taqqu (1986), "Weak convergence of moving averages with infinite variance," in: E. Eberlain and M.S. Taqqu (eds.), *Dependence in Probability and Statistics: A Survey of Recent Results.* Boston: Birkhäuser.

Avram, F. and M.S. Taqqu (1989a), "Probability bounds for M–Skorohod oscillations," *Stochastic Processes and their Applications* 33, 63–72.

Avram, F. and M.S. Taqqu (1989b), "Weak convergence of sums of moving averages in the α–stable domain of attraction," mimeo, Boston University.

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

Blattberg, R.C. and N.J. Gonedes (1974), "A comparison of the stable and Student distributions as statistical models for stock prices," *Journal of Business* 47, 244–280.

Bollerslev, T. (1986), "Generalized autoregressive conditional heteroskedasticity," *Journal of Econometrics* 31, 307–327.

Brockwell, P.J. and R.A. Davis (1987), *Time Series: Theory and Methods.* New York: Springer.

Brown, R.L., J. Durbin, and J.M. Evans (1975), "Techniques for testing the constancy of regression relationships over time," *Journal of the Royal Statistical Society, Series B* 37, 149–163.

Chambers, J.M., C.L. Mallows, and B.W. Stuck (1976), "A method for simulating stable random variables," *Journal of the American Statistical Association* 71, 340–344.

Davis, R.A. and S.I. Resnick (1985a), "Limit theory for moving averages of random variables with regularly varying tail probabilities," *Annals of Probability* 13, 179–195.

Davis, R.A. and S.I. Resnick (1985b), "More limit theory for sample correlation functions of moving averages," *Stochastic Processes and their Applications* 20, 257–279.

Davis, R.A. and S.I. Resnick (1986), "Limit theory for the sample covariance and correlation functions of moving averages," *Annals of Statistics* 14, 533–558.

Durbin, J. (1969), "Tests for serial correlation in regression analysis based on the periodogram of least squares residuals," *Biometrika* 56, 1–15.

Dumouchel, W.H. (1983), "Estimating the stable index α in order to measure the tail thickness: A critique," *Annals of Statistics* 11, 1019–1031.

Engle, R.F. (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation," *Econometrica* 50, 987–1008.

Engle, R.F. and T. Bollerslev (1986), "Modeling the persistence of conditional variances," *Econometric Reviews* 5, 1–50.

Fama, E.F. (1965), "The behavior of stock market prices," *Journal of Business* 38, 34–105.

Fielitz, B.D. and J.P. Rozelle (1982), "Stable distributions and mixtures of distributions hypotheses for common stock returns," *Journal of the American Statistical Association* 78, 28–36.

Hall, J.A., B.W. Brorsen, and S.H. Irwin (1989), "The distribution of futures prices: A test of the stable–Paretian and mixture of normals hypothesis," *Journal of Financial and Quantitative Analysis* 24, 105–116.

Hall, P. (1982), "On some simple estimates of an exponent of regular variation," *Journal of the Royal Statistical Society, Series B* 44, 37–42.

Hall, P. and C.C. Heyde (1980), *Martingale Limit Theory and its Application.* New York: Academic Press.

Hamilton, J.D. (1988), "Rational–expectations econometric analysis of changes of regime: An investigation of the term structure of interest rates," *Journal of Economic Dynamics and Control* 12, 385–423.

Hamilton, J.D. (1989), "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica* 57, 357–384.

Hannan, E.J., and M. Kanter (1977), "Autoregressive processes with infinite variance," *Journal of Applied Probability* 14, 411–415.

Haubrich, J.G., and A.W. Lo (1989), "The sources and nature of long–term memory in the business cycle," mimeo, Wharton School, University of Pennsylvania.

Hill, B.M. (1975), "A simple general approach to inference about the tail of a distribution," *Annals of Mathematical Statistics* 3, 1163–1174.

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

Kanter, M. and W.L. Steiger (1974), "Regression and autoregression with infinite variance," *Advances in Applied Probability* 6, 768–783.

Logan, B.F., C.L. Mallows, S.O. Rice, and L.A. Shepp (1973), "Limit distributions of self-normalized sums," *Annals of Probability* 1, 788–809.

Mandelbrot, B.B. (1963), "The variation of certain speculative prices," *Journal of Business* 36, 394–419.

Newey, W.K., and K.D. West (1987), "A simple, positive, semi–definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica* 55, 703–708.

Pagan, A.R., and G.W. Schwert (1989a), "Testing for covariance stationarity in stock market data," mimeo, University of Rochester. (Forthcoming, *Economics Letters.*)

Pagan, A.R., and G.W. Schwert (1989b), "Alternative models for conditional stock volatility," mimeo, University of Rochester. (Forthcoming, *Journal of Econometrics.*)

Phillips, P.C.B. (1987), "Time series regression with a unit root," *Econometrica* 55, 277–301.

Phillips, P.C.B. (1990), "Time series regression with a unit root and infinite variance errors," *Econometric Theory* 6, 44–62.

Phillips, P.C.B., and V.A. Hajivassiliou (1987), "Bimodal t–ratios," Cowles Foundation Discussion Paper No. 842, July.

Phillips, P.C.B., and V. Solo (1989), "Asymptotics for linear processes," Cowles Foundation Discussion Paper No. 932, October.

Ploberger, W., and W. Krämer (1986), "On studentizing a test for structural change," *Economics Letters* 20, 341–344.

Resnick, S.I. (1986), "Point processes regular variation and weak convergence," *Advances in Applied Probability* 18, 66–138.

Schwert, G.W. (1989), "Business cycles, financial crises, and stock volatility," *Carnegie–Rochester Conference Series* 31, 83–129.

So, J. (1987), "The sub–Gaussian distribution of currency futures: stable Paretian or nonstationary?," *Review of Economics and Statistics* 69, 100–107.