Cowles Foundation Paper 866

Journal of Empirical Finance 1 (1994) 211–248. North-Holland

Testing the covariance stationarity of heavy-tailed time series

An overview of the theory with applications to several financial datasets

Mico Loretan*

University of Wisconsin, Madison, WI, USA

Peter C.B. Phillips*

Yale University, New Haven, CT, USA

Final version accepted December 1992

Abstract

We investigate methods of testing the proposition that the unconditional variance of a time series is constant over time. Motivated by the observation that many financial datasets are "heavy-tailed," we focus on the properties of statistical tests of covariance stationarity when unconditional fourth and second moments of the data are not finite. We find that sample split prediction tests and cusum of squares tests have nonstandard limiting distributions when fourth unconditional moments are infinite. These tests are consistent provided that variances are finite. However, the rate of divergence under the alternative hypothesis and hence the power of these tests is sensitive to the index of tail thickness in the data. We estimate the maximal moment exponent (which measures tail thickness) for a number of stock market return and exchange rate return series, and conclude that fourth unconditional moments of these series do not appear to be finite. In our formal tests of covariance stationarity, we reject the null hypothesis of constancy of the unconditional variance of these series. This raises questions about the nature of the observed volatility in economic time series, and about appropriate methods of statistically modeling this volatility.

1. Introduction

An interesting feature of many empirical time series of stock market and exchange rate returns data is the apparent volatility or nonstationarity in their

Correspondence to: Dr. P.C.B. Phillips, Yale University, Department of Economics, New Haven, CT 06520, USA.

*We are most grateful to Adrian Pagan and Bill Schwert for supplying their stock return data used in Pagan and Schwert (1990a). Our thanks further go to the editor and two anonymous referees for constructive comments, as well as to the NSF for research support to Phillips under Grant No. SES 8821180. This paper constitutes an extensive revision of parts of our earlier paper (Phillips and Loretan (1990)).

SSDI 0927-5398(93)E0005-M

sample variances. Motivated by the observed serial dependence in volatility patterns, it has become common practice in finance to model conditional variances as functions of past conditional variances and past squared innovations, say as ARCH and GARCH processes (Engle (1982), Bollerslev (1986)). However, these models do impose restrictions on the properties of *unconditional* moments of the data. Stationary GARCH series, for instance, require constancy of the unconditional first and second moments: and IGARCH series (Engle and Bollerslev (1986)) allow only for special types of drift of the unconditional second moments (Nelson (1990)).

Constancy of the unconditional second moments of a time series is rarely implied by models of optimizing behavior of economic agents. Indeed, the efficient markets hypothesis is typically formulated either in terms of restrictions on the conditional mean of the data – e.g., expected excess returns on an asset given only public information should be zero – or in terms of relationship between conditional mean and conditional variance, as in the CAPM model. One may reasonably expect that unconditional second moments would *not* be constant over long periods of time. The speed at which information reaches traders, their ability to interpret this information, and the availability of sophisticated tools such as computerized trading mechanisms are all subject to temporal evolution and can be hypothesized to affect the *unconditional* variance of financial assets such as stocks, stock market aggregates, and foreign exchange.

Notwithstanding these reasonable empirical expectations, the assumption of covariance stationarity is convenient in time series analysis, and it is an assumption that is frequently employed, whether explicitly or implicitly, in much applied research. Indeed, covariance stationarity is often assumed more for statistical convenience rather than for good economic reasons. In the case of wide-sense stationary ARCH and GARCH models it is the byproduct of the specification of the dynamics of the conditional second moments, as we remarked above. Models of "switching regimes" (Hamilton (1988, 1989)) assume that each economic regime is characterized, *inter alia*, by a different variance of the relevant variables: integrating over all possible regimes, one finds that this class of models also implies constancy of the unconditional variances. Given the role of covariance stationarity in these important lines of empirical research, it would seem that testing the maintained hypothesis of covariance stationarity in empirical time series analysis is important in itself and relevant to the debate how best to model the volatility that is observed in the markets for common stock and foreign exchange.

Pagan and Schwert (1990a, 1990b) have recently presented some persuasive evidence that the unconditional variance of U.S. stock market returns data cannot be assumed to be constant over long periods of time. The testing methodology of Pagan and Schwert (1990a) relies on an auxiliary maintained assumption that fourth unconditional moments of the data are finite. This raises the natural question whether their rejection of covariance stationarity is caused by a failure of this auxiliary assumption; i.e., whether their finding is merely the byproduct of a "thick tail" phenomenon in the data generating process, or is indeed due to failure of covariance stationarity. Most asymptotic distribution theory used in econometric research relies on moment conditions which carefully control outlier occurrences. It is not unusual in time series analysis to see conditions of the type "let all required moments exist." However, in financial and commodity market time series the extent of outlier activity casts doubt on the suitability of such generic moment conditions. Mandelbrot (1963) provided suggestive evidence that even second moments may not exist for this type of data, and he proposed stable distributions with infinite variance as an alternative to finite-variance statistical models. Subsequent research (Blattberg and Gonedes (1974), Fielitz and Rozelle (1982). Akgiray and Booth (1989)) has generally reached the conclusion that second moments of most datasets appear to be finite. The existence of higher-order, especially fourth, moments has been studied less extensively. As we shall show below, the empirical evidence from the monthly stock market return data used by Pagan and Schwert (1990a, 1990b) suggests that fourth moments are infinite; and hence, their maintained hypothesis is not supported by their data. However, their conclusion that U.S. stock market returns are not covariance stationary over long periods of time is strongly supported even under a more general maintained hypothesis.

In this paper, we provide an asymptotic theory of tests of covariance stationarity which are based on estimated sample variances. Specifically, we examine properties of moments-based tests of the null of constancy of the unconditional variance; our methods could easily be extended to study constancy of unconditional autocovariances as well. We consider a sample split prediction test, the cusum of squares test, and the rescaled range test. All of these tests were employed in the Pagan–Schwert studies cited above. While our focus is on tests of covariance stationarity, we remark that there are other applications of our theory, for instance, to LM tests for ARCH and GARCH effects which also routinely rely on fourth moment conditions.

The paper is organized as follows. Section 2 presents an overview of the theory of various tests of covariance stationarity for heavy-tailed time series. Moment condition failure is introduced explicitly through the assumption that the tails of the innovation distribution are of the asymptotic Pareto–Lévy type. The conventional theory for sample moment based tests of covariance stationarity, which applies when fourth moments are finite, involves standard-normal and Brownian bridge asymptotics. When fourth moments fail we show that the new limit theory involves asymmetric stable distributions. The correct critical values in these situations are typically smaller than in the standard case, so that tests based on normal asymptotics are conservative. We further show that when unconditional second moments are infinite the tests are inconsistent, i.e., that they do not diverge under the alternative of changing dispersion over time. When second moments are finite the tests are consistent but diverge at a slower rate than $\sqrt{n}$. The tests are robust to serial dependence and conditional heterogeneity in the data generating process.

In Section 3 we propose a method of estimating the *maximal moment exponent* (i.e., the maximal degree to which moments exist) of a time series directly. Our proposed estimator is simple to compute and possesses a standard-normal limiting distribution. In Section 4 we summarize the results of Monte Carlo experiments which we conducted to obtain the critical values of moment based tests of covariance stationarity. In Section 5 we consider the empirical properties of monthly and daily U.S. stock market return series and of several daily exchange rate return series. Estimates of the maximal moment exponents of these series strongly indicate that fourth moments are infinite but that second moments are finite. In order to test covariance stationarity of these series, critical values based on the new asymptotic theory in Section 2 must be employed. Using this theory, we reject the null that unconditional variances are constant over time for all the series studied. Section 6 offers some concluding comments and suggestions for further research.

2. Overview of the theory of testing the covariance stationarity of heavy-tailed time series

To make this paper self-contained we present here a brief overview of the theory of testing for covariance stationarity, and focus on the consequences of failure of fourth moment conditions for the resulting limit theory. The results given here were derived in an earlier paper by the authors (Phillips and Loretan (1990)), which is available on request. The first subsection below defines terminology and states the asymptotic convergence results that we need; the second subsection gives the limit theory for several types of tests of covariance stationarity.

2.1. Preliminaries

Let $\{\varepsilon_t\}_{t=1}^{n}$ be an iid sequence of innovations whose tail behavior is of the asymptotic Pareto–Lévy form, *viz.*

$$\begin{aligned} \Pr(\varepsilon > x) &= pC^{\alpha}x^{-\alpha}(1+\alpha_1(x)), \quad x>0, \\ \Pr(\varepsilon > -x) &= qC^{\alpha}x^{-\alpha}(1+\alpha_2(x)), \quad x>0, \end{aligned} \tag{C1}$$

where $\alpha_i(x) \to 0$ $(i=1,2)$ as $x \to \infty$. The symmetry parameters p and q satisfy $p \geqslant 0$, $q \geqslant 0$, $p+q=1$. The parameter $C>0$ is a scale dispersion parameter.

The most important parameter in determining the tail shape in (C1) is α, which is the *maximal moment exponent* of the distribution. Absolute moments of ε of order less than α are finite, while all higher-order moments are infinite, i.e., $\alpha = \sup\{a > 0 : \mathrm{E}|\varepsilon|^a < \infty\}$. When $0 < \alpha < 2$, condition (C1) ensures that ε lies in the *normal domain of attraction* of a stable law with *characteristic exponent* α, and we shall write $\varepsilon \in \mathcal{N}D(\alpha)$ to signify this fact. When $\alpha > 2$, ε is in the domain of attraction of a normal distribution, so that standardized partial sums of ε converge in distribution to a normal distribution. When $2 < \alpha < 4$ it is important to note that $\varepsilon^2 \in \mathcal{N}D(\alpha/2)$ and $\alpha/2 < 2$, so that partial sums of ε^2, properly standardized, no longer converge weakly to a normal distribution. This will be of relevance for tests which do not rely on finite fourth moments.

We add the following centering condition, which applies when $\alpha \geqslant 1$:

$$\textit{If } \alpha > 1 \textit{ then we require } \mathrm{E}(\varepsilon) = 0. \textit{ If } \alpha = 1 \textit{ we require } \varepsilon =_d -\varepsilon \\ (\textit{i.e., } \varepsilon \textit{ is distributed symmetrically about the origin}). \tag{C2}$$

No centering is required when $\alpha < 1$.

Suppose that the observed time series is generated by the linear process

$$y_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}\,, \tag{1}$$

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

$$\sum_{j=1}^{\infty} j|c_j|^p < \infty \quad \text{for} \quad 0 < p < \alpha,\ p \leqslant 1\,. \tag{C3}$$

Note that (C3) holds whenever y_t is generated by a stationary ARMA process because then the coefficients in equation (1) decline geometrically and thus trivially satisfy (C3) (see, e.g., Brockwell and Davis (1991)). In what follows it will be convenient for us to explicitly work with the AR(p) process

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t\,, \tag{2}$$

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

Assumption (C3) is convenient for the development of the asymptotic theory given in the following subsection. While a wide class of time series are covered by (1) and (C3), the main limiting requirement as far as financial time series applications are concerned is the use of iid innovations in (1). This assumption is not as restrictive as it may appear. As shown in Loretan (1991. ch. 1) for the case of finite fourth moments the same limit theory as that given below applies also when the innovations are martingale differences, which allows the data to be conditionally heterogeneous and weakly dependent under the null. The results stated below therefore apply to a fairly wide class of stationary time series.

Under (C1) and (C2) with $0 < \alpha < 2$, the appropriate scaling factor for partial sums of ε is given by $a_n = Cn^{1/\alpha}$. Note that when $\alpha = 2$ this leads to conventional $\sqrt{n}$ asymptotics. We shall employ the following limit theory (where r, $0 < r \leqslant 1$, represents a certain fraction of the overall sample):

$$a_n^{-1} \sum_{t=1}^{n} \varepsilon_t \rightarrow_d U_\alpha(1)\,, \tag{3}$$

$$a_n^{-1} \sum_{t=1}^{[nr]} \varepsilon_t \rightarrow_d U_\alpha(r)\,, \tag{4}$$

$$\left(a_n^{-1} \sum_{t=1}^{[nr]} \varepsilon_t,\, a_n^{-2} \sum_{t=1}^{[nr]} \varepsilon_t^2 \right) \rightarrow_d \left(U_\alpha(r), \int_0^r (dU_\alpha)^2 \right). \tag{5}$$

Here $U_\alpha(r)$ is the Lévy α-stable process on $D[0, 1]$, the space of CADLAG functions on the $[0, 1]$ interval, and $\int_0^r (dU_\alpha)^2 = [U_\alpha]_r$ is its quadratic variation process. See Ibragimov and Linnik (1971, ch. 2), Resnick (1986) and Phillips (1990) for further details on these concepts and results. The symbol "$\rightarrow_d$" denotes weak convergence on $D[0, 1]$ of the associated probability measures (Billingsley (1968)). The symbol "$\rightarrow_{fdd}$" will be used below when only the finite-dimensional distributions converge, i.e., when (4) and (5) only apply for a finite number of values of r (taken jointly) rather than on the function space $D[0, 1]$ itself.

When $2 < \alpha < 4$ we have $n^{-1}\sum_1^n \varepsilon_t \rightarrow_p 0$ and $n^{-1}\sum_1^n \varepsilon_t^2 \rightarrow_p \sigma_\varepsilon^2$. But since $\varepsilon^2 \in \mathscr{N}D(\alpha/2)$ we also have a stable limit distribution theory for the centered sample second moments. We find

$$a_n^{-2} \sum_{t=1}^{n} (\varepsilon_t^2 - \sigma_\varepsilon^2) \rightarrow_d U_{\alpha/2}(1)\,, \tag{6}$$

$$a_n^{-2} \sum_{t=1}^{[nr]} (\varepsilon_t^2 - \sigma_\varepsilon^2) \rightarrow_d U_{\alpha/2}(r)\,, \tag{7}$$

$$\left(a_n^{-2} \sum_{t=1}^{[nr]} (\varepsilon_t^2 - \sigma_\varepsilon^2),\quad a_n^{-4} \sum_{t=1}^{[nr]} (\varepsilon_t^2 - \sigma_\varepsilon^2)^2 \right) \rightarrow_d \left(U_{\alpha/2}(r), \int_0^r (dU_{\alpha/2})^2 \right). \tag{8}$$

The stable process $U_{\alpha/2}(r)$ in equation (7) is asymmetric with maximal positive skewness. Finally, we observe that if y_t is generated by the AR(p) process (2) the coefficients are estimated consistently by the OLS regression

$$y_t = \sum_{i=1}^{p} \hat{\phi}_i y_{t-i} + \hat{\varepsilon}_t , \tag{9}$$

irrespective of the value of α (see Kanter and Steiger (1974) and Hannan and Kanter (1977)). Correspondingly, the OLS residual $\hat{\varepsilon}_t$ is consistent for ε_t for all $\alpha > 0$. We may therefore use these residuals in place of ε_t in the limit theory above.

2.2. *Asymptotic theory for tests of covariance stationarity*

There are several ways of testing for homogeneous unconditional variances. Those we shall consider are based on the behavior of the sample second moments of either the observed time series y_t or of the residual series $\hat{\varepsilon}_t$ which are obtained from a consistent autoregression as in equation (9).

When testing for constancy of the unconditional variances over time, it is natural to start by splitting the sample into two eras according to $n = n_1 + n_2$ with $n_1 = k_n n_2$ and to consider the null hypothesis

$$\mathrm{H}_0: \ \mathrm{E}\hat{\mu}_2^{(1)} = \mathrm{E}\hat{\mu}_2^{(2)}, \qquad \text{where}$$

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

(We shall use the affixes "$^{(1)}$" and "$^{(2)}$" throughout this paper to signify quantities that pertain to the first and second subera, respectively.) Putting $\hat{\tau} = \hat{\mu}_2^{(1)} - \hat{\mu}_2^{(2)}$, the null hypothesis can be restated as $\mathrm{H}_0 : \mathrm{E}\hat{\tau} = 0$. To determine whether $\hat{\tau}$ is "significantly" different from zero we must estimate its variation. Define a kernel-based estimate of the "long-run" variance of the squared observations of y_t as

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

where $\hat{\gamma}_j$ is the j-th serial covariance of y_t^2 and l is a suitable lag truncation number. When $\alpha > 4$ this estimator of the long-run variance of y_t^2 is consistent provided that $l \to \infty$ as $n \to \infty$ in such a way that $l^2/n \to 0$ (see, e.g., Newey and West (1987) and Andrews (1991)).

Suppose that $k_n \to k > 0$. Then $k/(1+k)$ (respectively, $1/(1+k)$) is the fraction of the overall sample in the limit that is spent in the first (second) era. The limiting sample variance of $\hat{\tau}$ is therefore given by $(1+k)v^2$, and we have for $\alpha > 4$ that $n^{1/2}\hat{\tau} \to_d N(0,(1+k)v^2)$. We now define the *sample split prediction test* as the standardized statistic

$$V_k(\tau) = ((1+k_n)\hat{v}^2)^{-1/2}\, n_1^{1/2}\hat{\tau}.$$

The limit theory for this test depends critically on whether $\alpha > 4$ or $\alpha < 4$. Specifically, we have:

Proposition 1 (Phillips and Loretan (1990), Theorem 3.1). Assume that (C1)–(C3) *hold and that $k_n \to k > 0$ (fixed) as $n \to \infty$. Then*

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

$$V_k(\tau) \to_d \left(k\int_0^1 (dU_{\alpha/2})^2\right)^{-1/2} \times ((1+k)U_{\alpha/2}(k/(1+k)) - kU_{\alpha/2}(1)) = \bar{V}_k,\ \textit{say}. \tag{10}$$

The random element $U_{\alpha/2}(r)$ is a maximally asymmetric stable process with characteristic exponent $\alpha/2$.

- *If $0 < \alpha < 4$ and $k = 1$,*

$$V_1(\tau) \to_d \left(\int_0^1 (dU^s_{\alpha/2})^2\right)^{-1/2} \times U^s_{\alpha/2}(1)\,. \tag{11}$$

The random element $U^s_{\alpha/2}(r)$ is a symmetric stable process with characteristic exponent $\alpha/2$.

Remarks: (i) When $0 < \alpha < 4$ the limit distribution of the sample split prediction test is described by the variate $\bar{V}_k$, which is a ratio of dependent stable variates, (ii) Comparing equations (10) and (11), the choice of $k = 1$ leads to a substantial simplification of the limiting distribution. Phillips and Loretan (1990) further show that the expression in (11) reduces to a standard normal law when $\alpha > 4$. (iii) The limiting distributions given by (10) and (11) are the same as the limit distribution of a "self-normalized sum" or t-ratio statistic (see Logan *et al.* (1973)). We provide plots of its density for the cases of $\alpha/2 = 1.5$ and $\alpha/2 = 1.05$ in Figures 1a and 1b.[1] A striking feature of these limit laws is their bimodality,

[1]Methods of obtaining estimates of the density and associated quantiles by Monte Carlo simulation are discussed in Section 4.

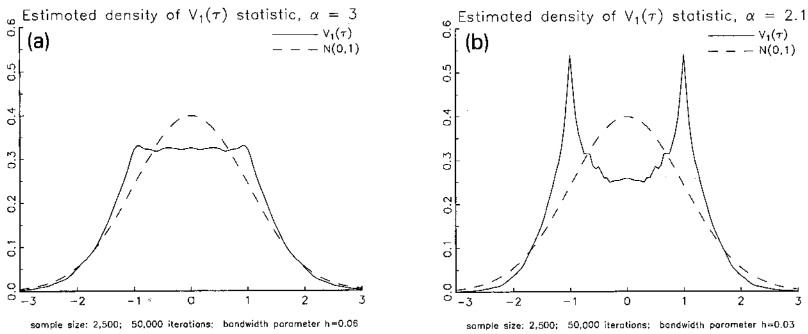

Fig. 1.

which becomes more pronounced as $\alpha \downarrow 2$. (iv) The same limit distribution for $V_k(\tau)$ obtains when the long-run variance of y_t^2, i.e., v^2, is estimated as a weighted average of the long-run variances in the respective suberas, i.e., as $\tilde{v}^2 = \hat{v}^{2^{(1)}} + k_n \hat{v}^{2^{(2)}}$. (v) Finally, the use of the consistent residuals $\hat{\varepsilon}_t$ in place of y_t leads to the same limiting distribution of $V_k(\tau)$ as that stated in Proposition 1. This means that the properties of the test are not affected by applying a linear filter such as (9) to the data.

Under the alternative hypothesis, measures of the unconditional dispersion are unequal in the two subperiods. For the finite variance case ($\alpha > 2$) we may express the alternative as

$$H_1: \quad \sigma_\varepsilon^{2^{(1)}} \neq \sigma_\varepsilon^{2^{(2)}}.$$

For the infinite variance case ($\alpha < 2$) we introduce heterogeneity through the scale dispersion coefficient C in (C1):

$$H_1': \quad C^{(1)} \neq C^{(2)}.$$

The following theorem from Phillips and Loretan (1990) describes the consistency properties of the sample split prediction test $V_k(\tau)$.

Proposition 2 (Phillips and Loretan (1990), Theorem 3.5). Assume that (C1)–(C3) *hold and* $k > 0$. *Two cases apply:*

$\alpha > 2$ *Under* H_1, *the statistic* $V_k(\tau)$ *diverges as* $n \to \infty$. *The rate of divergence is given by*

$$V_k(\tau) = O_p(n^{1/2}) \quad \textit{for} \quad \alpha > 4 \tag{12}$$

$$V_k(\tau) = O_p(n^{1-2/\alpha}) \quad \textit{for} \quad 2 < \alpha \leqslant 4. \tag{13}$$

$0 < \alpha \leqslant 2$ *Under* H'_1 *the statistic* $V_k(\tau)$ *is inconsistent. Specifically, as* $n \to \infty$

$$V_k(\tau) = O_p(1)\,. \tag{14}$$

These results also hold when the estimator $\hat{v}^2$ *is replaced by* $\tilde{v}^2$ *or if the consistent residuals* $\hat{\varepsilon}_t$ *are used in place of* y_t.

The inconsistency of the sample split prediction test in the infinite variance case is not surprising: the test relies on a comparison of sample second moments across suberas – when variances are infinite, the sample second moments are naturally poor measures of the population dispersion. The intermediate case ($2 < \alpha \leqslant 4$) is of particular interest: The test is consistent, but its rate of divergence is slower than in the standard ($\alpha > 4$) case. The presence of large outliers in both suberas therefore makes it less likely that "significant" differences between the subera estimates of μ_2 will be found in finite samples. We shall refer to this finding again in Section 5 below, where we need to interpret the findings from an application of this test to time series of stock market and exchange rate returns.

The *cumulative sum (cusum) of squares test* offers an alternative way of testing the null of covariance stationarity. The test is based on the cumulative sums of $y_t^2 - \hat{\mu}_2$, where $\hat{\mu}_2 = n^{-1} \sum_1^n y_t^2$ is the unconditional mean of y_t^2, and is defined as

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

This test is a studentized version of the cusum of squares test originally proposed by Brown, Durbin and Evans (1975), since it standardizes the partial sums of $(y_t^2 - \hat{\mu}_2)$ by a sample-based estimate of v^2 rather than by its expected value under normality. Ploberger and Krämer (1986) originally analyzed this modification. Pagan and Schwert (1990a) also use this form of the cusum of squares statistic. Using the consistent residuals $\hat{\varepsilon}_t$ instead of y_t, we define the analogous statistic

$$\psi_n^{\varepsilon}(r) = (n\hat{v}_{\varepsilon}^2)^{-1/2} \sum_{t=1}^{[nr]} (\hat{\varepsilon}_t^2 - \hat{\sigma}_{\varepsilon}^2)\,,$$

where $\hat{\sigma}_{\varepsilon}^2 = n^{-1}\sum_1^n \hat{\varepsilon}_t^2$ is the sample variance of the residuals and $\hat{v}_{\varepsilon}^2$ is a kernel-type estimate of the long-run fourth moment of $\hat{\varepsilon}_t$ defined analogously to $\hat{v}^2$.

Since $k = r/(1-r)$ and $r = k/(1+k)$, we note the following correspondence between the sample split prediction test $V_k(\tau)$ and the cusum of squares test $\psi_n(r)$

for *fixed* $r \in [0, 1]$ and $k > 0$:

$$V_{r/(1-r)}(\tau) = (r(1-r))^{1/2} \cdot \psi_n(r),$$

$$\psi_n(k/(1+k)) = (k^{1/2}/(1+k))\, V_k(\tau).$$

For $\alpha > 4$ and $n \to \infty$, the cusum of squares statistic converges weakly to a Brownian bridge (a tied-down Brownian motion or Wiener process) on $C[0, 1]$, the space of continuous functions on the interval $[0, 1]$ (see Billingsley (1968)). The limit law is again quite different when $\alpha < 4$. We give the following result:

Proposition 3 (Phillips and Loretan (1990), Theorem 3.6). Assume that (C1)–(C3) *hold. Then*

- *If* $\alpha > 4$

$$\psi_n(r),\ \psi_n^{\varepsilon}(r) \to_d W(r) - r \cdot W(1) = B(r), \tag{15}$$

a standard Brownian Bridge on $C[0, 1]$, $W(r)$ *is a standard Wiener process on* $C[0, 1]$.

- *If* $\alpha < 4$

$$\psi_n(r) \to_{fdd} K_{\alpha/2}(r) \Big/ \left(\int_0^1 (dU_{\alpha/2})^2 \right)^{1/2} = L_{\alpha/2}(r), \text{ say} \tag{16}$$

$$\psi_n^{\varepsilon}(r) \to_d L_{\alpha/2}(r), \tag{17}$$

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

Remarks: (i) The limit distribution of the cusum of squares statistic is given by a standardized tied-down stable-Lévy process when $\alpha < 4$; this is the natural generalization of the Brownian bridge result for the case of finite fourth moments. (ii) For $\alpha < 4$ weak convergence only obtains for the statistic $\psi_n^{\varepsilon}(r)$ but not for $\psi_n(r)$. Serial dependence in y_t under the null can lead to complications in function space convergence, and only the finite-dimensional distributions of $\psi_n(r)$ converge in general. (See Avram and Taqqu (1992) for a discussion of this difficulty.) Note that the problem does not arise with $\psi_n^{\varepsilon}(r)$. For this cusum statistic we have both finite dimensional and function space convergence.

Inference may be conducted in several ways. First, one may use the finite-dimensional distributions (fdd's) of the cusum of squares, i.e, consider the behavior of the statistic for *fixed* values of r, and compare the test values against the applicable critical values (these are computed in Section 4). This procedure generalizes the use of the sample split prediction test discussed above. It suffers from the disadvantage that inferences may be sensitive to the choice of r: the finite-dimensional distributions of $\psi_n(r)$ may lie outside the confidence bands for some values of r but not for others, leading to inconclusive outcomes.

Instead, one may proceed by considering the trajectory of the cusum of squares over the whole sample, i.e., for all values of r, and derive the probabilities that the test will exceed any given value *somewhere* on the $[0, 1]$ interval. For this type of analysis, one needs to work with the statistic $\psi_n^\varepsilon(r)$, since the statistic $\psi_n(r)$ does not converge weakly on $D[0, 1]$ in general when $\alpha < 4$ as discussed above. One may construct scalar-valued test statistics such as $\sup_r(\psi_n^\varepsilon(r))$, $\sup_r|\psi_n^\varepsilon(r)|$, or $\sup_r(\psi_n^\varepsilon(r)) - \inf_r(\psi_n^\varepsilon(r))$. The second of these statistics is also known as the Kolmogorov–Smirnov test, and the third is often referred to as the rescaled range test or "R/S statistic" (Hurst (1951), Mandelbrot (1972), Lo (1991)). From Proposition 3 and the continuous mapping theorem, we find immediately that

$$\sup_r(\psi_n^\varepsilon(r)) \to_d \sup_r L_{\alpha/2}(r)\,,$$

$$\sup_r|\psi_n^\varepsilon(r)| \to_d \sup_r|L_{\alpha/2}(r)|\,,\ \text{and}$$

$$\sup_r(\psi_n^\varepsilon(r)) - \inf_r(\psi_n^\varepsilon(r)) \to_d \sup_r L_{\alpha/2}(r) - \inf_r L_{\alpha/2}(r) = R_L,\ \text{say}\,.$$

Critical values of these tests are also provided in Section 4 below.

The consistency properties of the cusum of squares test $\psi_n^\varepsilon(r)$ may be studied in the same way as the sample split prediction tests. The rates of divergence of the test are comparable to those given in Proposition 2 above. The cusum of squares test has decreasing power as $\alpha \downarrow 2$ and is inconsistent for $0 < \alpha \leqslant 2$:

Proposition 4 (Phillips and Loretan (1990), Theorem 3.7). *Suppose* (C1)–(C2) *and* (1) *hold. Then we have*

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

$$\psi_n^\varepsilon(r) = O_p(n^{1/2}), \quad \textit{for} \quad \alpha > 4\,, \tag{18}$$

$$\psi_n^\varepsilon(r) = O_p(n^{1-2/\alpha}), \quad \textit{for} \quad 2 < \alpha \leqslant 4\,. \tag{19}$$

$0 < \alpha \leqslant 2$ *Under* H'_1 *tests based on* $\psi_n^\varepsilon(r)$ *are inconsistent and we have as* $n \to \infty$

$$\psi_n^\varepsilon(r) = O_p(1)\,. \tag{20}$$

We close this section by pointing out that these propositions do not depend critically on the assumption that the innovations are iid or that the observables are generated by a linear process from these innovations. The propositions can be extended to allow for conditional heteroskedasticity and more general weakly dependent sequences of random variables under the null (Loretan (1991) provides some extensions along these lines). What matters for purposes of testing constancy of the *unconditional* variance is the maximal finite moment of the data, denoted by α here, and in particular whether it exceeds 4, lies in the interval $(2, 4]$, or is less than or equal to 2.

3. Direct estimates of maximal moment exponents

The preceding section demonstrates that it is important to assess tail shape characteristics if we are to understand the properties of tests such as those of covariance stationarity in practice. Estimating tail shapes is of course an empirical issue. One way to proceed is to attempt to estimate the maximal moment exponent directly from the data. A convenient and easy to implement method of estimating the parameters α and C in the Pareto–Lévy model (C1) is available and proceeds as follows. Let $\{\hat{\varepsilon}_t\}_{t=1}^{n}$ be the residuals from (9) and let $\hat{\varepsilon}_{n1} \leqslant \hat{\varepsilon}_{n2} \leqslant \cdots \leqslant \hat{\varepsilon}_{nn}$ be the order statistics corresponding to this sample of residuals. Define the estimators

$$\hat{\alpha}_s = \left(s^{-1} \sum_{j=1}^{s} \ln \hat{\varepsilon}_{n,n-j+1} - \ln \hat{\varepsilon}_{n,n-s} \right)^{-1}, \tag{21}$$

$$\hat{C}_s = s \cdot n^{-1} (\hat{\varepsilon}_{n,n-s})^{\hat{\alpha}_s}, \tag{22}$$

for some integer s. It is assumed that n is large enough and s/n small enough so that $\hat{\varepsilon}_{n,n-s} > 0$ and thus $\hat{\alpha}_s$ and $\hat{C}_s$ are well defined real quantities. These estimators were originally proposed by Hill (1975) as conditional maximum likelihood estimators of the maximal moment exponent α and the scale dispersion coefficient C in (C1). The asymptotic theory for these estimators in the general case of a distribution whose tails have the asymptotic Pareto–Lévy form (C1) is due to Hall (1982), who shows that it is optimal, at least in terms of asymptotic bias and variance of these estimators, 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 (C1). (If the tails of the distribution are exactly Pareto we may set $\gamma = \infty$.)

There is some advantage in 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)

$$s^{1/2}(\hat{\alpha}_s - \alpha) \to_d N(0, \alpha^2), \quad \text{and} \tag{23}$$

$$s^{1/2}(\ln(n/s))^{-1}(\hat{C}_s - C) \to_d N(0, C^2). \tag{24}$$

These asymptotics have the advantage that the limit distributions (23) and (24) involve only scale nuisance parameters which are easily eliminated in statistical tests. Notice that these estimates pertain to the right or upper tail of the distribution of ε; to estimate the parameters of the left or lower tail, one simply multiplies the order statistics by -1 and repeats the computations.

These formulae provide a method of estimating the parameters α and C and of computing their asymptotic standard errors. They also lead to the following statistical tests of the hypothesis that the parameters are constant across subperiods. Consider the null hypotheses

$$\mathrm{H}_{0,\alpha}\colon\quad \alpha^{(1)} = \alpha^{(2)} = \alpha \quad \text{and} \quad \mathrm{H}_{0,C}\colon\quad C^{(1)} = C^{(2)} = C\,.$$

Suppose that we split the sample such that $n_1 = n_2 = n/2$. Let $\hat{\alpha}_s^{(i)}$ and $\hat{C}_s^{(i)}$, $i = 1, 2$ denote the subsample estimates of α and C, and put $\hat{\tau}_{\alpha,s} = \hat{\alpha}_s^{(1)} - \hat{\alpha}_s^{(2)}$ and $\hat{\tau}_{C,s} = \hat{C}_s^{(1)} - \hat{C}_s^{(2)}$. Define the sample split prediction tests

$$\tilde{V}(\tau_\alpha) = s^{1/2}\cdot \hat{\tau}_{\alpha,s}/(\hat{\alpha}_s^{(1)2} + \hat{\alpha}_s^{(2)2})^{1/2}\,, \qquad \text{and} \tag{25}$$

$$\tilde{V}(\tau_C) = s^{1/2}\cdot(\ln(n/s))^{-1}\cdot \hat{\tau}_{C,s}/(\hat{C}_s^{(1)2} + \hat{C}_s^{(2)2})^{1/2}\,. \tag{26}$$

We have the following limit theory for these tests under the null and under the alternative hypotheses

$$\mathrm{H}_{1,\alpha}\colon\quad \alpha^{(1)} \neq \alpha^{(2)} \quad \text{and} \quad \mathrm{H}_1'\colon\quad C^{(1)} \neq C^{(2)}\,.$$

Proposition 5 (Phillips and Loretan (1990), Theorems 3.8–3.10). Let (C1), (C2) *and* (1) *hold and suppose that* $k_n \to k = 1$ *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* $s/n^{2\gamma/(2\gamma+\alpha)} \to 0$ as $n_1 \to \infty$. *Then*

$$\tilde{V}(\tau_\alpha) \to_d N(0,1) \text{ under } \mathrm{H}_{0,\alpha}\,, \tag{27}$$

$$\tilde{V}(\tau_C) \to_d N(0,1) \text{ under } \mathrm{H}_{0,C}\,, \tag{28}$$

$$\tilde{V}(\tau_\alpha) = O_p(s^{1/2}) \text{ under } \mathrm{H}_{1,\alpha}\,, \tag{29}$$

$$\tilde{V}(\tau_C) = O_p(s^{1/2}/\ln(n_1/s)) \text{ under } \mathrm{H}_1'\,. \tag{30}$$

The limit distribution of both sample split prediction tests is thus very simple. They are consistent irrespective of the value of α and would thus appear to enjoy certain advantages over the moment based tests $V_k(\tau)$ and $\psi_n^\varepsilon(r)$ discussed above. The cost of this consistency is that their rate of divergence is potentially quite slow, given the dependence of s on the secondary tail shape parameter γ. The tests should therefore not be employed to test for structural breaks when fourth moment condition failure is not a feature of the data. (Notice, however, that if the tails are exactly Paretian ($\gamma = \infty$), the rate of divergence of both tests is $O_p(n^{1/2})$.)

In finite samples, the choice of the nuisance parameter s, i.e., the number of order statistics included in the computations, may affect the point estimates and standard errors, and hence the inferences drawn from the data. Loretan (1991, ch. 3) reports the results from an extensive simulation study which was designed to assess the influence of the choice of s on statistical inference. He considered values of $\alpha = \{0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 10, 100\}$, sample sizes $n = \{1{,}000, 4{,}000, 10{,}000\}$ and $s = \{100, 125, 150, 175, 200\}$ for the case of $n_1 = n/2 = 500$. (He let s grow more slowly than n in the other two cases.) We summarize his results as follows:

- The integer s should generally not exceed about $0.1 \cdot n_1$; larger choices of s typically involve including more than just tail observations in equations (21) and (22), which severely biases point estimates of α and C.
- The estimator (21) has good centering and size properties for all values of α that were considered.
- The estimator (22) has good centering properties but its simulated variance is larger than its asymptotic variance (C^2) for all but the largest sample sizes.
- The test $\tilde{V}(\tau_\alpha)$ has very good empirical size properties under the null, i.e., its actual rejection frequency is close to the nominal size of the test given asymptotic $N(0, 1)$ critical values. This finding does not seem to depend on the precise value of α.
- The test $\tilde{V}(\tau_C)$, on the other hand, tends to overreject the corresponding null hypothesis (H_1') for all nominal sizes; this size distortion decreases only very slowly when larger sample sizes are considered. However, this tendency can be corrected by heuristically adjusting the required nominal size.
- The empirical power of the tests, even for small variations of α and C across the suberas, was satisfactory in the simulations. This seems to be due in part to the large sample sizes that were employed.

In any case, and especially when the precise distributional properties of the tails of the data are unknown in an empirical application, it is advisable to estimate α for a variety of trial values of s. This is also the procedure suggested by Dumouchel (1983).

In Section 5 below we shall see that the unconditional distribution of the tails or "outlier observations" of many stock market and exchange rate return series are very well described as being of the Pareto–Lévy type. Direct estimates of the maximal moment exponents of the series yield point estimates which are less than 4 and greater than 2, which indicates that variances of the data are finite but also that there is still enough outlier activity to cast into doubt the existence of fourth moments in the population.

4. Monte Carlo simulations of tests of covariances stationarity

In Section 2 the limiting distributions of three types of tests of covariance stationarity were given for all $\alpha > 0$. For $\alpha < 4$ the limit laws of these statistics are given by functionals of stable processes. Closed form expressions for these distributions are unknown, so that 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 need to 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)[2] for the general asymmetric case. We consider values of $\alpha = \{2.1, 2.5, 3.0, 3.5, 3.8\}$, and set $n = 1{,}000$ as the "large" sample size, except for the asymmetric case when $\alpha = 2.1$, where we found that it was necessary to set $n = 2{,}500$ to approximate the asymptotic distributions adequately. We performed 50,000 repetitions of all experiments. To increase efficiency, the symmetry and skew-symmetry of the distributions were exploited in computing critical values and densities. All simulations and estimations were carried out in the GAUSS programming language.[3]

We first study the large sample distribution of the sample split prediction test statistic $V_1(\tau)$. We performed the following simulation experiment. For fixed α, draw n iid *symmetric* stable random numbers x_i of index $\alpha/2$, set $y_i = n^{-2/\alpha} \cdot x_i$ – note that by self-similarity $y_i \approx_d dU^s_{\alpha/2}$ – and compute the ratio $(\sum_i^n y_i)/(\sum_i^n y_i^2)^{1/2}$, thereby simulating the quantity $U^s_{\alpha/2}(1)/(\int_0^1 (dU^s_{\alpha/2})^2)^{1/2}$. Critical values, at the usual levels of significance, are shown in Table 1, together with the standard normal critical values which apply when $\alpha > 4$. The new critical values for typical test sizes are all lower than the conventional ones. For a two-sided test of 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 parameter α is less than 4 leads to conservative tests.

Nonparametric kernel estimates of the density of the $V_1(\tau)$ statistic, for the cases of $\alpha = 3$ and $\alpha = 2.1$, were computed from the simulations and are graphed in Figures 1a and 1b. These densities are quite different from the standard

[2] We used the formula (2.3) in Chambers *et al.* (1976). We did not use the algorithm proposed in their equation (4.1), which is based on a modified skewness parameter β', since we are only interested in generating *maximally skewed* stable variates, and (2.3) is faster to compute than (4.1). Note that when $1 < \alpha/2 < 2$ one must set the skewness parameter $\beta = -1$ in (2.3) to obtain stable variates with maximal *positive* skewness.

[3] Copies of the programs are available on request.

Table 1

Critical values of the sample split prediction test statistic $V_1(\tau)$.

$\Pr(X < c)$	α 2.1	2.5	3.0	3.5	3.8	$N(0, 1)$ $[\alpha > 4]$
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

Remarks: 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 respective null hypotheses.

normal density. They are *bimodal* and *platykurtic*: the modes are located at -1 and $+1$, and the tails are thinner than that of the standard normal pdf. This reflects the fact that their critical values at the 1% and 5% levels are smaller than those of the standard normal distribution. 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.[4]

We turn to the empirical distribution of the limit law $L_{\alpha/2}(r)$ of the cusum of squares statistics $\psi_n(r)$ and $\psi_n^{\varepsilon}(r)$. To this effect, we draw n iid *asymmetric* stable variates x_i of index $\alpha/2$, set $y_i = n^{-2/\alpha} \cdot x_i$, and compute $(\sum_i^{[nr]} y_i - r \cdot \sum_i^n y_i)/(\sum_i^n y_i^2)^{1/2}$ as the large sample approximation to $(U_{\alpha/2}(r) - r \cdot U_{\alpha/2}(1))/(\int_0^1 (dU_{\alpha/2})^2)^{1/2} = L_{\alpha/2}(r)$, 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's) of $L_{\alpha/2}(r)$ are given in Tables 2a–2e; the exact critical values for the fdd's of the Brownian Bridge process $B(r)$ are shown in Table 2f for comparison. Only the upper confidence contours are provided; the lower confidence contours are obtained from the skew-symmetric relationship $L_{\alpha/2}(r) =_d - L_{\alpha/2}(1 - r)$. In Figures 2a–2f 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 about zero in the vertical axis (Table 2f and Figure 2f), the contours become increasingly asymmetric as $\alpha \downarrow 2$. In comparison to the standard case of $\alpha > 4$, for $\alpha < 4$ the upper confidence contours increase faster when r is close to 0,

[4] Logan *et al.* (1973) further show that the density of the t-ratio statistic is finite and continuous if $1 < \alpha/2 < 2$, but has poles at -1 and $+1$ if $\alpha/2 < 1$.

Table 2

Critical values of the finite dimensional distributions of the cusum of squares test statistics $\psi_n(r)$ and $\psi_n^{\varepsilon}(r)$.

	r								
Pr($X < c$)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
(2a) $\alpha = 2.1$									
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
(2b) $\alpha = 2.5$									
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
(2c) $\alpha = 3.0$									
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
(2d) $\alpha = 3.5$									
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
(2e) $\alpha = 3.8$									
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
(2f) $\alpha > 4.0$									
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

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

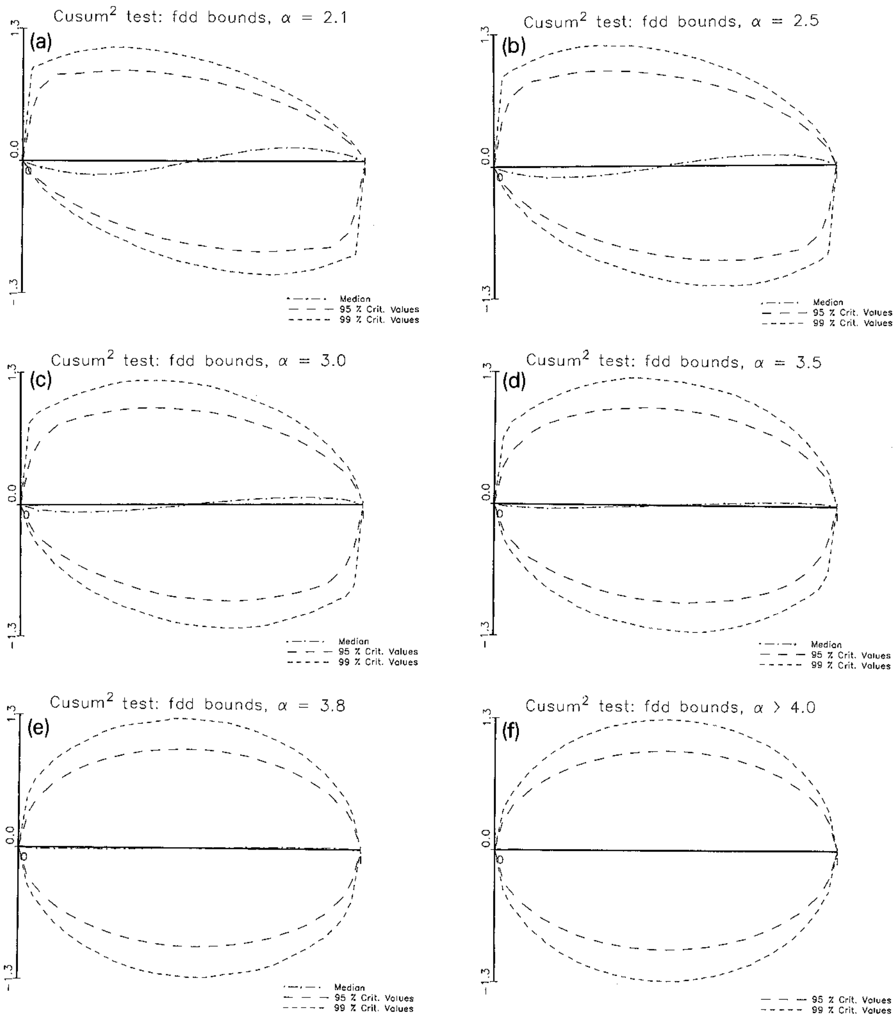

Fig. 2.

attain a smaller maximal height, and returns to 0 more slowly as r approaches 1. Second, the medians of the fdd's of $L_{\alpha/2}(r)$ also depend on r: the medians are negative for $r < 0.5$ and positive for $r > 0.5$. The fdd's of $L_{\alpha/2}(r)$, for $\alpha = \{2.1, 3.0, 4.0\}$, are further contrasted in Figures 3a and 3b, in which we graph their 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 $r \geqslant 0.5$, but become increasingly liberal as $r \downarrow 0$. E.g., for $r = 0.5$, the (2-sided) 99% critical value decreases from 1.29 (for $\alpha > 4$) to 1.19

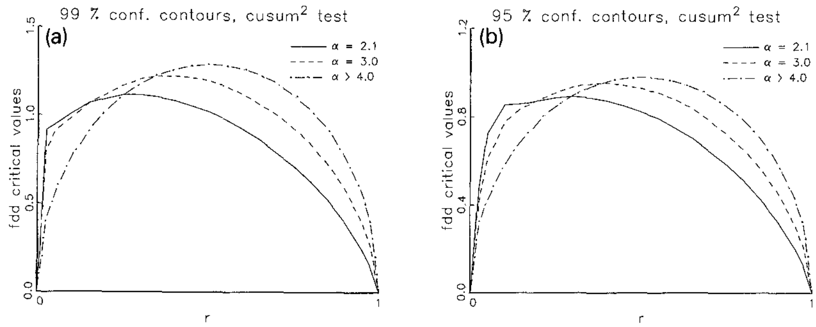

Fig. 3.

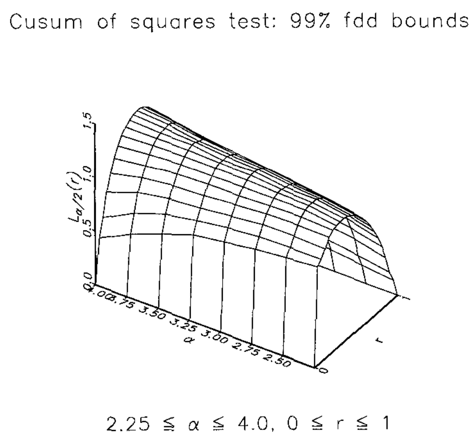

Fig. 4.

($\alpha = 3$) and further to 0.96 ($\alpha = 2.1$): use of the conventional critical value leads to conservative inference 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 4 summarizes the dependence of the shape of the confidence contours of the fdd's on α in a three-dimensional graph, for $2.25 \leqslant \alpha \leqslant 4$ and $0 \leqslant r \leqslant 1$.

To complement the information on the fdd critical values given in Tables 2a–2e, we provide the asymptotic critical values for the statistic $\sup_r(\psi_n^\varepsilon(r))$, delivered from simulating $\sup_r(L_{\alpha/2}(r))$, in Table 3. The critical values of the statistic $\inf_r(\psi_n^\varepsilon(r))$ are obtained by applying the skew-symmetric relationship $\Pr(\inf_r(L_{\alpha/2}(r) < -c) = \Pr(\sup_r(L_{\alpha/2}(r) > c)$, $c > 0$. As discussed in Section 2, use of these critical values does not lead to the size distortions that may result from considering only the fdd's 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* as α decreases from 4 to 2, 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 trajectory 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 applications 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 rescaled range statistic $R_L = \sup_r(\psi_n^\varepsilon(r)) - \inf_r(\psi_n^\varepsilon(r))$ are given in Table 4, together with critical values of the conventional ($\alpha > 4$) range statistic R_B. As was the case for the statistic $\sup_r(L_{\alpha/2}(r))$ above, critical values of R_L decrease as α declines from 4 to 2, so that $\alpha > 4$-based

Table 3

Critical values of the supremum statistic $\sup_r(\psi_n^\varepsilon(r))$.

	α					
$\Pr(X < c)$	2.1	2.5	3.0	3.5	3.8	$N(0, 1)$ $[\alpha > 4]$
80%	0.67	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

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

Table 4

Critical values of the range statistic R_L.

	α					
$\Pr(X < c)$	2.1	2.5	3.0	3.5	3.8	$N(0, 1)$ $[\alpha > 4]$
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	1.863
99%	1.48	1.72	1.80	1.87	1.93	2.001

Remarks: For critical values are based on 50,000 simulations of the test statistics for $\alpha < 4$, with a sample size of $n = 1{,}000$ (except in the case of $\alpha = 2.1 : n = 2{,}500$). For $\alpha > 4$, the exact critical values solve the equation $\Pr(R_B < c) = 1 + 2\sum_{k=1}^{\infty}((1 - 4k^2c^2)\cdot\exp(-2k^2c^2))$ (Lo (1991), equation (3.9)).

critical values are again conservative when the true value of α is less than 4. For a one-sided test of size 5%, say, the appropriate critical value decreases from 1.75 ($\alpha > 4$) to 1.57 ($\alpha = 3$) and further to 1.31 ($\alpha = 2.1$).

5. Empirical application: Some properties of stock market and exchange rate returns series

5.1. The data

In this section, we analyze the outlier behavior and the null of covariance stationarity for two stock market return series and five exchange rate return series. The two stock market series are (i) monthly returns to a broad index of U.S. stocks[5] from January 1834 to December 1987 ($n = 1{,}848$) and (ii) daily returns to the "Standard and Poors 500" stock market index, from July 1962 to December 1987 ($n = 6{,}405$); the latter data were obtained from the 1988 CRSP tape. The exchange rate returns are based on daily closing spot market prices measured relative to the U.S. dollar. The following countries were selected: France, Germany, Japan, Switzerland, and the United Kingdom.[6] The exchange rate series were extracted from the EHRA dataset compiled by the Board of Governors of the Federal Reserve System, and span the period from December 1978 to January 1991. The number of observations is 3,140 in all cases except for Japan, where we have $n = 3{,}134$.[7]

We are interested in answering the following questions: (i) What are the point estimates of the maximal moment exponent α of these seven time series, and is there evidence of fourth moment or even 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 constant over time? (iii) Can we support the finding of Pagan and Schwert (1990a) 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 infinite?

Since our statistical theory assumes that the series to be tested for variance constancy have zero mean, we must estimate and filter out the conditional mean of our datasets before we can perform the tests. Thus, prior to the analysis of the data presented below, all series were demeaned as follows: we regressed the data

[5] The construction of this dataset is discussed in Schwert (1989).

[6] These five series are a subset of the 18 exchange rate series analyzed by Loretan (1991, ch. 3). The present selection is made purely in order to concentrate on the properties of some "major" and frequently studied exchange rate series. The conclusions reached from using the other exchange rate series are quite similar to the ones presented here.

[7] This minor difference in number of observations is due to a different coincidence of national holidays between the United States and Japan on the one hand and the United States and the European countries on the other.

on five weekday dummies (daily-frequency data only) and (in a second step) on twelve monthly dummies to eliminate potential day-of-week and seasonal effects, and passed the residuals from these regressions through a long AR filter to remove any linear serial dependence.[8] The residuals from these transformations are used in the computations reported below. We tested for and found evidence of serial dependence in the second moments of the residuals. However, these pretests did not indicate that the fitted parameters from a GARCH model would lie on the IGARCH boundary. Since the estimated amount of serial dependence in the higher moments did not rule out covariance stationarity, we did not eliminate it from the residuals. As discussed in Section 2, the use of the long run variance estimator of v^2 leads to test statistics which are robust to serial dependence in the conditional second moments under the null and alternative.

5.2. Point estimates of maximal moment exponents

We first study the tail behavior of our series informally, by plotting the tails of their empirical distribution functions in double-logarithmic coordinates. (More precisely, we graph $\log_{10} x$ against $\Pr(X > \log_{10} x)$, for $x > 0$.) In these coordinates, Pareto-like tails of the distributions form straight lines with a slope equal to $-\alpha$. As can be seen from Figures 5a–5g, about 5% to 10% of the observations fall into the right tail of the corresponding distribution. In all cases the "outliers" of the respective series seem to be very well described as being distributed according to a Pareto law. (The left tails of the data, while not presented here, also appear to follow a Pareto law.)

We formally estimate the maximal moment exponents using formula (21) above. We computed α for a variety of choices of s, the number of included order statistics. Following the Monte Carlo-based evidence presented in Section 3 above and the suggestions by Dumouchel (1983), s was set so as not to exceed ten percent of the sample size. We report the point estimates of α in Tables 5a–5g for the seven series, separately for the right and left tails. The associated asymptotic standard errors are delivered from equation (23) above, and were computed under the assumption that $s = o(n^{2\gamma/(2\gamma+\alpha)})$, for some $\gamma > 0$. (The standard errors could be sharpened considerably if we assume that the tails were exactly Pareto and thus assumed $s = o(n)$.)

The point estimates are almost all less than 4; they range from 2.5 to 3.2 for the monthly stock market return series, from 3.1 to 3.8 for the daily stock market return series,[9] and from 2.4 to 3.7 for most of the exchange rate return series.

[8] A detailed description of all preliminary data transformation is given in Phillips and Loretan (1990) for the stock market series and in Loretan (1991, ch. 3) for the exchange rate series.

[9] Since the daily and monthly stock index series cover different time horizons, and are defined differently, we would not expect their respective point estimates of α to be close.

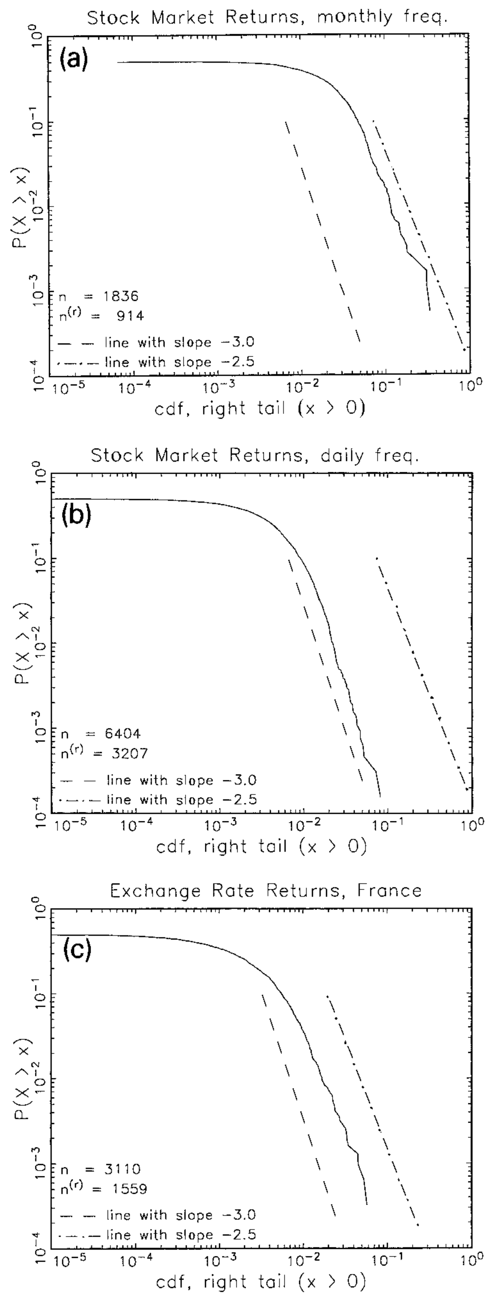
Stock Market Returns, monthly freq.
(a)
P(X > x)
cdf, right tail (x > 0)
n = 1836
n^(r) = 914
line with slope −3.0
line with slope −2.5
Stock Market Returns, daily freq.
(b)
P(X > x)
cdf, right tail (x > 0)
n = 6404
n^(r) = 3207
line with slope −3.0
line with slope −2.5
Exchange Rate Returns, France
(c)
P(X > x)
cdf, right tail (x > 0)
n = 3110
n^(r) = 1559
line with slope −3.0
line with slope −2.5

Fig. 5(a–c).

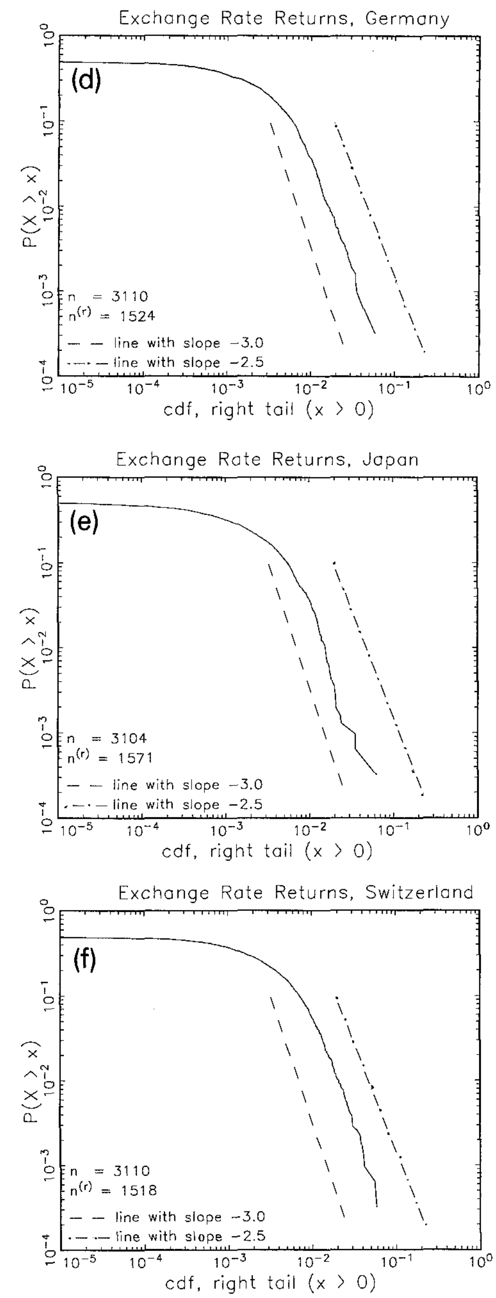
Exchange Rate Returns, Germany
(d)
n = 3110
$n^{(r)}$ = 1524
line with slope −3.0
line with slope −2.5
P(X > x)
cdf, right tail (x > 0)
Exchange Rate Returns, Japan
(e)
n = 3104
$n^{(r)}$ = 1571
line with slope −3.0
line with slope −2.5
P(X > x)
cdf, right tail (x > 0)
Exchange Rate Returns, Switzerland
(f)
n = 3110
$n^{(r)}$ = 1518
line with slope −3.0
line with slope −2.5
P(X > x)
cdf, right tail (x > 0)

Fig. 5(d–f).

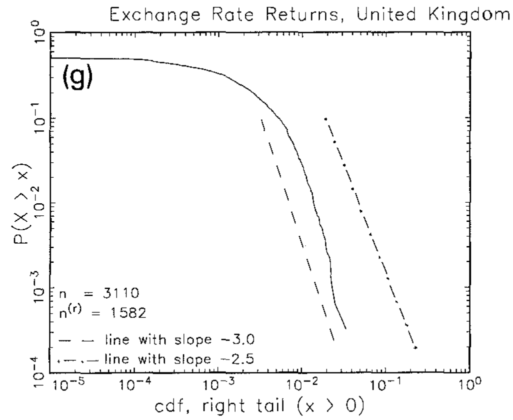

Fig. 5(g).

(The only exceptions occur for Switzerland at $s = 20$ in the left tail and for the United Kingdom at $s = 20$ in the right tail (cf. Tables 5f and 5g).) Notice that for several of these series the point estimates of α are remarkably insensitive to the choice of s: e.g., for the right tail of the German exchange rate returns (Table 5d) the point estimates fluctuate between 2.97 ($s = 20$) and 2.79 ($s = 100$). Except in a few cases – which only occur for small values of s, i.e., when the standard errors are very large – these point estimates are all more than two asymptotic standard deviations below 4.

The point estimates of α are also greater than 2, implying that infinite variance is not a feature of these datasets. While the tails of the two empirical distributions are certainly 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 law with a characteristic exponent $\alpha < 2$. These direct estimates of the maximal moment exponent of the distributions contribute new evidence that relates to the long standing debate as to whether to model stock returns as realizations of stable laws. Mandelbrot's (1963) seminal work studied the behavior of 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 prices and returns, do not follow stable laws and are better modeled by finite variance

Table 5

Point estimates of maximal moment exponents.

(5a) Monthly-frequency U.S. stock market returns

	left tail ($n^{(l)} = 992$)		right tail ($n^{(r)} = 914$)	
s	$\hat{\alpha}_s^{(l)}$	s.e.$_s^{(l)}$	$\hat{\alpha}_s^{(r)}$	s.e.$_s^{(r)}$
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)

(5b) Daily-frequency U.S. stock market returns

	left tail ($n^{(l)} = 3{,}197$)		right tail ($n^{(r)} = 3{,}207$)	
s	$\hat{\alpha}_s^{(l)}$	s.e.$_s^{(l)}$	$\hat{\alpha}_s^{(r)}$	s.e.$_s^{(r)}$
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)

(5c) Exchange rates: France

	left tail ($n^{(l)} = 1{,}551$)		right tail ($n^{(r)} = 1{,}559$)	
s	$\hat{\alpha}_s^{(l)}$	s.e.$_s^{(l)}$	$\hat{\alpha}_s^{(r)}$	s.e.$_s^{(r)}$
20	2.85	0.64	2.86	0.64
30	3.00	0.55	2.38	0.43
50	2.73	0.39	2.29	0.32
75	2.42	0.28	2.60	0.30
100	2.39	0.24	2.52	0.25

(5d) Exchange rates: Germany

	left tail ($n^{(l)} = 1{,}586$)		right tail ($n^{(r)} = 1{,}524$)	
s	$\hat{\alpha}_s^{(l)}$	s.e.$_s^{(l)}$	$\hat{\alpha}_s^{(r)}$	s.e.$_s^{(r)}$
20	3.11	0.70	2.97	0.66
30	3.70	0.68	2.90	0.53
50	3.44	0.49	2.70	0.38
75	3.03	0.35	2.80	0.32
100	3.16	0.32	2.79	0.28

Table 5. (continued)

(5e) Exchange rates: Japan

	left tail ($n^{(l)} = 1{,}533$)		right tail ($n^{(r)} = 1{,}571$)	
s	$\hat{\alpha}_s^{(l)}$	s.e.$_s^{(l)}$	$\hat{\alpha}_s^{(r)}$	s.e.$_s^{(r)}$
20	3.09	0.69	3.59	0.80
30	3.03	0.55	4.38	0.80
50	3.09	0.44	3.64	0.51
75	3.05	0.35	3.73	0.43
100	2.83	0.28	3.30	0.33

(5f) Exchange rates: Switzerland

	left tail ($n^{(l)} = 1{,}592$)		right tail ($n^{(r)} = 1{,}518$)	
s	$\hat{\alpha}_s^{(l)}$	s.e.$_s^{(l)}$	$\hat{\alpha}_s^{(r)}$	s.e.$_s^{(r)}$
20	5.13	1.15	3.30	0.74
30	3.69	0.67	3.03	0.55
50	3.41	0.48	2.97	0.42
75	3.35	0.39	2.71	0.31
100	3.10	0.31	2.77	0.28

(5g) Exchange rates: United Kingdom

	left tail ($n^{(l)} = 1{,}528$)		right tail ($n^{(r)} = 1{,}582$)	
s	$\hat{\alpha}_s^{(l)}$	s.e.$_s^{(l)}$	$\hat{\alpha}_s^{(r)}$	s.e.$_s^{(r)}$
20	3.25	0.73	4.71	1.05
30	3.00	0.55	3.92	0.72
50	2.98	0.42	3.56	0.50
75	2.89	0.33	3.44	0.40
100	2.59	0.26	3.27	0.33

distributions. The point estimates of α presented here agree with this general result, at least as far as aggregate series are concerned.

But the simple observation that the unconditional variances appear to be finite obviously does not suffice as a characterization of the distributions' tail behavior. In particular, it does not indicate the order of magnitude of the tails, or how many higher-order moments of the distributions can be assumed to exist. Since our point estimates of α are generally significantly less than 4, fourth moment condition failure is an important feature of the data and should affect the way tests of covariance stationarity are carried out for these time series.

We recognize that empirical distributions have finite support in practice and therefore have finite sampling 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* other distribution that has infinite support, including the normal distribution, and does not provide a framework for describing outlier activity present in the data. The "outliers" in all series we consider here, as evidenced by the form of the cdf's plotted in Figures 5a–5g, are rather well described as being distributed according to a Pareto law. Note that our plots too, by extrapolation, would assign a negligible probability of observing changes of the daily-frequency series of, say, plus or minus 100% per day: the Pareto-tail hypothesis is therefore not in conflict with economic common sense. To summarize, saying that the support of an empirical distribution is bounded says very little about the nature of outlier activity that may occur in the data. In contrast, a model of Paretian tails not only appears to provide an adequate model of observed outlier activity, but also gives a predictive framework for the rate at which outliers appear. Finally, it permits the development of an asymptotic theory of tests of covariance stationarity when outlier activity plays a significant role empirically.

To assess the robustness of the finding that the tails of stock returns and exchange rate returns are Pareto-like, we formally test the equality of point estimates of α across the right and left tail and across time periods, using the sample split prediction test statistic $\tilde{V}(\tau_\alpha)$ given in equation (25) above. From (27) the limit distribution of this test is standard normal, and the test is $O_p(\sqrt{n})$-consistent under the alternative if the tails are exactly Paretian. The test results are reported in Tables 6a–6g.[10] We cannot reject the null that the same parameter α applies to both the left and right tails of the distributions (cf. the first column of the Tables), and cannot reject constancy of α over time (cf. the second column) for any of the series except for the Japanese exchange rate series. For the latter series, we reject constancy of α in favor of the alternative that α has increased over time, i.e., that the relative importance of outliers has decreased. We interpret these findings as evidence in support of the assumption that α is a useful summary parameter in describing the empirical distributions of these series, and that any heterogeneity present in the data is therefore most likely to be due to variation in the dispersion parameters (C or σ_ε^2) alone.[11]

[10] To avoid clutter, we do not report the values of $\hat{\alpha}^{(i)}$, $i = 1, 2$, separately in the tables. The estimates of $\hat{\alpha}^{(i)}$ which enter the $\tilde{V}(\tau_\alpha)$ tests are reported in detail in Table 3(a) in Loretan (1991, ch. 3), and are available on request.

[11] Observe that when α changes, the variance of the series changes as well, unless an exactly offsetting change in C occurs. Our moments-based tests of covariance stationarity reported below are thus not affected directly by the issue of constancy of the tail shape parameter α. The only place where α matters for moments-based tests is when one must decide whether these tests are "statistically significant," because their critical values do depend on α when $\alpha < 4$.

Table 6

Sample split prediction test $\tilde{V}(\tau_\alpha)$ of constancy of the tail parameter α, across tails and over time.

(6a) Monthly-frequency U.S. 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

(6b) Daily-frequency U.S. 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

(6c) Exchange rate returns: France

s	Constancy of α across tails	s	Constancy of α over time
50	0.867	30	− 0.408
75	− 0.427	40	− 0.839
100	− 0.397	50	− 1.283

(6d) Exchange rate returns: Germany

s	Constancy of α across tails	s	Constancy of α over time
50	1.192	30	− 1.413
75	− 0.489	40	− 1.550
100	− 0.878	50	− 2.000*

(6e) Exchange rate returns: Japan

s	Constancy of α across tails	s	Constancy of α over time
50	− 0.821	30	− 2.303*
75	− 1.216	40	− 2.893**
100	− 1.076	50	− 3.529

(6f) Exchange rate returns: Switzerland

s	Constancy of α across tails	s	Constancy of α over time
50	0.692	30	− 0.883
75	1.297	40	− 1.699
100	0.807	50	− 2.680

(6g) Exchange rate returns: United Kingdom

s	Constancy of α across tails	s	Constancy of α over time
50	− 0.875	30	− 1.620
75	− 1.062	40	− 1.579
100	− 1.643	50	− 0.731

Recently, Jansen and de Vries (1991) have analyzed the tail shapes of stock returns for several U.S. companies, and have obtained results which are quite similar to those given here for our aggregate stock returns. Specifically, they find that (i) the tails or outliers of individual stock returns are well described by a Pareto law, (ii) constancy of the tail index α cannot be rejected for the series and sample periods they consider, and (iii) variances appear to be finite but the existence of fourth moments is more problematical. Our empirical work differs from theirs in the following two points. First, we use Hall's (1982) extensions to Hill's (1975) estimator of α because we do not wish to commit to a maintained hypothesis that the tails are *exactly* Paretian; this leads to larger asymptotic standard errors for our estimators. Second, we estimate α for a number of choices of s; Jansen and de Vries (1991) only appear to have performed their calculation for $s = 100$ (Table 2, p. 22). It is unclear from their discussion how sensitive their findings are to this choice of s. However, these differences in methodologies are minor, and our conclusions regarding tail shapes and moment condition failure for stock return series are remarkably similar.

5.3. Are stock returns and exchange rate returns covariance stationary?

Since all of our series appear to have finite variances but infinite fourth moments, moments-based tests of covariance stationarity are consistent but require the use of the new limit theory and the modified critical values presented in Sections 2 and 4 of this paper. Since serial dependence is present in the higher order moments of the data, it is important to choose the lag truncation parameter l (in the kernel estimator of v^2) large enough. After some informal data analysis, we set $l = 8$ for the monthly stock return series,[12] equal to 12 for the daily stock return series, and equal to 20 for the exchange rate series.[13]

Consider first the sample split prediction test $V_k(\tau)$ evaluated at $k = 1$, i.e., for suberas of equal length. (The behavior of the test for other values of k can be inferred from the trajectories of the cusum of squares tests discussed below.) The estimated values of this test statistic are given in Table 7. With the exception of the German exchange rate series, we reject the null of covariance stationarity for all of the series. Since all statistics are negative, the null is rejected in favor of *increased* unconditional variance from era 1 to era 2.

To assess how sensitive this conclusion may be to the choice of the break ratio, we now turn to the cusum of squares statistics for the same series. We plot the cusums in Figures 6a–6g, along with the corresponding 5% critical values of

[12] This is also the value used by Pagan and Schwert (1990a).

[13] A more sophisticated analysis would employ a formal data-based decision rule for the choice of the lag truncation parameter. However, it is unlikely that this modification would lead to substantially different results for the data considered here.

Table 7

Sample split prediction tests of covariance stationarity.

Series	Value of $V_1(\tau)$
Monthly Stock Returns	-3.34^{***}
Daily Stock Returns (S&P 500)	-2.06^{**}
Exchange Rates: France	-2.59^{***}
Exchange Rates: Germany	-1.52
Exchange Rates: Japan	-3.33^{***}
Exchange Rates: Switzerland	-3.04^{***}
Exchange Rates: U. Kingdom	-3.04^{***}

Remarks: Statistics significant at the 95% and 99% levels of confidence are highlighted with two and three asterisks, respectively. Critical values are obtained from Table 1.

fdd-based tests and of the infimum test.[14] Observe that the trajectories of the tests tend to lie outside the 95% fdd confidence contours for substantial portions of the sample periods, and that the infimum of the tests almost always exceeds the corresponding 95% critical value. The numerical values of the rescaled infimum and range test statistics for all series are provided in Table 8. We reject the null of covariance stationarity at the 99% level of confidence for all series on the basis of the infimum statistics, with the exception of the German exchange rate return series. Covariance stationarity is rejected for German exchange rate returns at the 90% level of confidence. The same conclusion follows from inspection of the range statistics. These results, as well as the ones presented earlier on the sample split prediction tests, are all the more notable since the tests for covariance stationarity, as shown in Section 2, have low power in finite samples against the alternative of changing variance in the presence of fourth moment condition failure.

Our empirical findings are thus twofold. First, all series considered here fail one or more of our tests of covariance stationarity. This throws into question the validity and robustness of studies that employ this assumption while analyzing 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 finding. However, the apparent stability of this tail shape parameter over long periods of time is an interesting empirical regularity or stylized fact that models of rational economic behavior should be designed to accommodate and explain.

[14]The plotted critical values are those which are appropriate when $\alpha > 4$. As discussed in Section 3, these values are actually conservative when $\alpha < 4$, i.e., the actual size of the test will be smaller than the nominal size (here: 5%). Thus, if the test exceeds the plotted critical values, rejection of the null is certainly warranted.

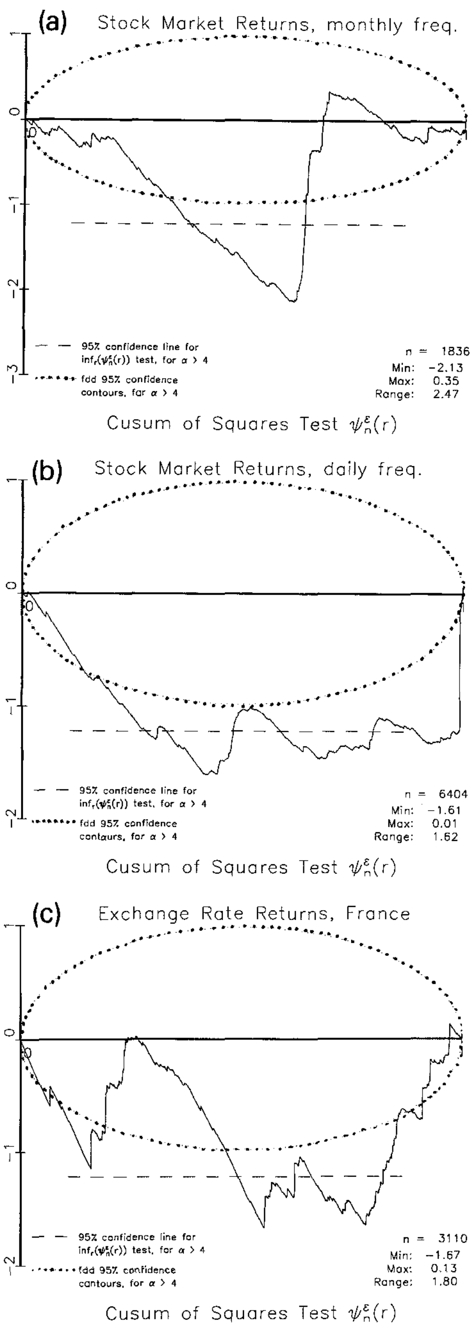
(a)
Stock Market Returns, monthly freq.
1
0
-1
-2
-3
95% confidence line for inf$_r$(ψ_n^ε(r)) test, for α > 4
fdd 95% confidence contours, for α > 4
n = 1836
Min: -2.13
Max: 0.35
Range: 2.47
Cusum of Squares Test ψ_n^ε(r)
(b)
Stock Market Returns, daily freq.
1
0
-1
-2
95% confidence line for inf$_r$(ψ_n^ε(r)) test, for α > 4
fdd 95% confidence contours, for α > 4
n = 6404
Min: -1.61
Max: 0.01
Range: 1.62
Cusum of Squares Test ψ_n^ε(r)
(c)
Exchange Rate Returns, France
1
0
-1
-2
95% confidence line for inf$_r$(ψ_n^ε(r)) test, for α > 4
fdd 95% confidence contours, for α > 4
n = 3110
Min: -1.67
Max: 0.13
Range: 1.80
Cusum of Squares Test ψ_n^ε(r)

Fig. 6(a–c).

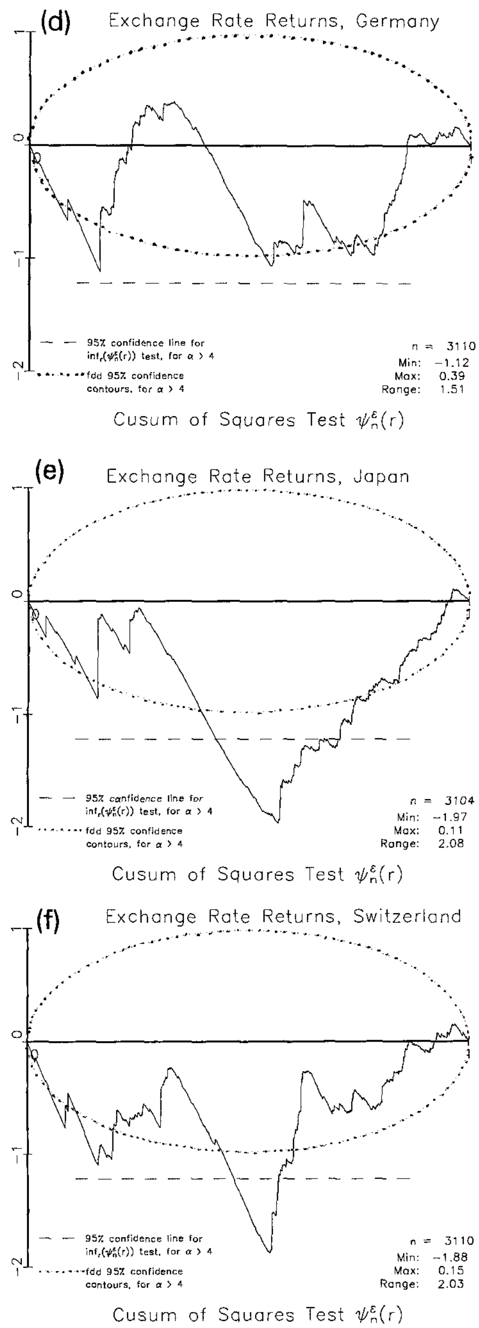
(d)
Exchange Rate Returns, Germany
95% confidence line for
fdd 95% confidence
contours, for α > 4
n = 3110
Min: −1.12
Max: 0.39
Range: 1.51
Cusum of Squares Test
(e)
Exchange Rate Returns, Japan
95% confidence line for
fdd 95% confidence
contours, for α > 4
n = 3104
Min: −1.97
Max: 0.11
Range: 2.08
Cusum of Squares Test
(f)
Exchange Rate Returns, Switzerland
95% confidence line for
fdd 95% confidence
contours, for α > 4
n = 3110
Min: −1.88
Max: 0.15
Range: 2.03
Cusum of Squares Test

Fig. 6(d–f).

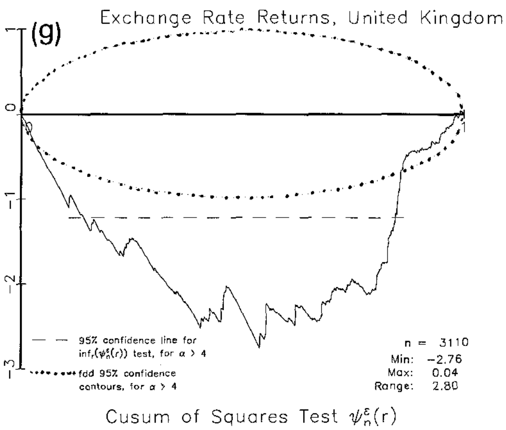

Fig. 6(g).

Table 8

Cusum of squares tests of covariance stationarity.

Series	Infimum statistic $\inf_r \psi_n^\varepsilon(r)$	Range statistic $\sup_r \psi_n^\varepsilon(r) - \inf_r \psi_n^\varepsilon(r)$
Monthly Stock Returns	− 2.13***	2.47***
Daily Stock Returns (S&P 500)	− 1.61***	1.62*
Exchange Rates: France	− 1.67***	1.80***
Exchange Rates: Germany	− 1.12*	1.51*
Exchange Rates: Japan	− 1.97***	2.08***
Exchange Rates: Switzerland	− 1.88***	2.03***
Exchange Rates: U. Kingdom	− 2.76***	2.80***

Remarks: Statistics significant at the 90%, 95% and 99% levels of confidence are highlighted with one, two and three asterisks, respectively. Critical values are obtained from Tables 3 and 4.

6. Conclusion

This paper presents 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. The tests are consistent as long as second moments are finite, but have low power against the alternative of unconditional heteroskedasticity when the maximal moment exponent of the distribution is close to 2. Critical values for these tests are obtained through Monte Carlo simulation.

Moments-based tests of variance constancy are thus insensitive or robust to moment condition failure in a fairly limited way: while standard critical values are incorrect, namely conservative, under the null when $\alpha < 4$, the tests remain consistent against the alternative of changing variance as long as $\alpha > 2$. There has been a movement in econometrics in recent years to develop and apply robust methods of specification or diagnostic testing. Here, robustness refers to the fact that the asymptotic properties of "robust tests" do not depend on potentially strong auxiliary assumptions under the null; this distinguishes such tests from standard LM tests of model misspecification (see, e.g., Godfrey (1988)). Instead, they require only much weaker regularity conditions. Wooldridge (1990, 1991) presents robust conditional moment tests for a wide variety of misspecifications of regression functions, and lists further references on this subject. A challenging topic for future research is to investigate methods of making tests of variance constancy "robust" by requiring weaker regularity conditions, while simultaneously allowing for fourth moment condition failure under the null.

We also provide methods of estimating the parameters α and C of distributions with Paretian tails, and show how to construct simple tests of the null that these parameters are constant over time. These tests have particularly simple, *viz.* standard normal, asymptotics.

In an empirical application, we test whether several stock market and exchange rate series have covariance stationary returns. No series pass the tests, an empirical finding that confirms earlier work by Pagan and Schwert (1990a) and casts doubt on the validity and descriptive accuracy of econometric models that assume the unconditional variances of these series to be constant. The tail shape parameter α of the empirical distributions is found to be constant across tails and over time for almost all series, indicating that even when fourth moments conditions and covariance stationarity fail there are still interesting empirical regularities in the series to be modeled.

We close by mentioning some possible extensions to our empirical analysis. Much work beckons: Given that covariance stationarity fails over long periods of time, how short must the intervals be chosen in order to apply models which assume unconditional homoskedasticity? Can one detect discrete break points in constant variance models? Do individual stocks behave similarly to the aggregate stock market series we have considered here? Finally, which theoretical models of rational economic behavior can plausibly explain and predict the apparent constancy of the tail shape and outlier activity of the data? All these issues seem worthy of further research.

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., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, Econometrica 59, 817–858.

Avram, F. and M.S. Taqqu, 1992, Weak convergence of sums of moving averages in the α-stable domain of attraction. Annals of Probability 20, 483–503.

Billingsley, P., 1968, Convergence of probability measures (Wiley, New York).

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

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

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 Society 71, 340–344.

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 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 stable distributions hypotheses for common stock returns, Journal of the American Statistical Association 78, 28–36.

Godfrey, L.G., 1988, Misspecification tests in econometrics. The Lagrange multiplier principle and other approaches (Cambridge University Press, Cambridge).

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.

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.

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

Hurst, H.E., 1951, Long-term storage capacity of reservoirs, Transactions of the American Society of Civil Engineers 116, 770–808.

Ibragimov, I.A. and Yu. V. Linnik, 1971, Independent and stationary sequences of random variables (Wolters-Noordhof, Groningen).

Jansen, D.W. and C.G. de Vries, 1991, On the frequency of large stock returns: Putting booms and busts into perspective, Review of Economics and Statistics 73, 18–24.

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

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

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.

Loretan, M., 1991, Testing Covariance Stationarity of Heavy-Tailed Economic Time Series. Unpublished Ph.D. Dissertation, Department of Economics, Yale University.

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

Mandelbrot, B.B., 1972, Statistical methodology for non-periodic cycles: From the covariance to R/S analysis, Annals of Economic and Social Measurement 1, 259–290.

Nelson, D.B., 1990, Stationarity and persistence in the GARCH(1.1) model, Econometric Theory, 6, 318–334.

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, 1990a, Testing for covariance stationarity in stock market data, Economics Letters 33, 165–170.

Pagan, A.R. and G.W. Schwert, 1990b, Alternative models for conditional stock volatility, Journal of Econometrics 45, 267–290.

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 M. Loretan, 1990, Testing covariance stationarity under moment condition failure with an application to common stock returns, Cowles Foundation Discussion Paper No. 947, June.

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.

Wooldridge, J.M., 1990, A unified approach to robust, regression-based specification tests, Econometric Theory 6, 17–43.

Wooldridge, J.M., 1991, On the application of robust, regression-based diagnostics to models of conditional means and conditional variances, Journal of Econometrics 47, 5–46.