

**STRUCTURAL CHANGE TESTS IN TAIL BEHAVIOUR
AND THE ASIAN CRISIS**

BY

**CARMELA QUINTOS,
ZHENHONG FAN
AND
PETER C.B. PHILLIPS**

COWLES FOUNDATION PAPER NO. 1030



**COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY**

Box 208281
New Haven, Connecticut 06520-8281
2002

Structural Change Tests in Tail Behaviour and the Asian Crisis

CARMELA QUINTOS
New York University

ZHENHONG FAN
New York University

and

PETER C. B. PHILLIPS
Yale University, University of Auckland and University of York

First version received January 1999; final version accepted August 2000 (Eds.)

This paper explores tests of the hypothesis that the tail thickness of a distribution is constant over time. Using Hill's conditional maximum likelihood estimator for the tail index of a distribution, tests of tail shape constancy are constructed that allow for an unknown breakpoint. The recursive test is shown to be inconsistent in one direction, and only a one-sided test is recommended. Specifically, the test can be used when the alternative hypothesis is that the tail index decreases over time. A rolling and sequential version of the test is consistent in both directions. The methods are illustrated on recent stock price data for Thailand, Malaysia and Indonesia. The period covers the recent Asian financial crisis and enables us to assess whether breakpoints in domestic asset return distributions are related to known changes in institutional arrangements in the foreign currency markets of these countries.

1. INTRODUCTION

The role of extreme outliers has long been relevant in economics and is now important in many financial market applications. More specifically, tail shape is well recognized as a distinguishing empirical feature of asset price and financial return data. Exchange rate and stock market return data are especially known to exhibit extreme outlier behaviour (Mandelbrot, 1963; Boothe-Glassman, 1987). Events like the October 1987 stock market crash, the Asian financial crisis of 1997–1998, the hedge fund crisis in 1998, and general market concern over value at risk (VaR) underscore the relevance of outlier activity and bear witness to the importance of this area of study in empirical work. One aspect of tail behaviour that is especially interesting during periods of market turbulence is whether the tail shape of the distribution itself changes, thereby increasing (or decreasing) the probability of outliers. The Asian financial crisis over 1997–1998 is a poignant recent example that we will consider, where abrupt changes in the market caught informed observers by surprise and were of a sufficient magnitude to suggest the possibility of a change in the underlying distribution.

There is econometric evidence that the tail behaviour of some financial series has changed over time. Assuming that the breakdate is exogenous or *known*, Phillips-Loretan

(1990) and Koedjick *et al.* (1990) construct tests for the null hypothesis that the tail thickness is constant over time. They find that the null of constancy is rejected by exchange rate data for the U.S., Japan and a subset of Western European countries. Other papers, like Pagan-Schwert (1990*a, b*), apply parameter constancy tests with stricter moment requirements (*i.e.* that the fourth moment be finite) and arrive at the same conclusion. Thus, there is a consensus from past empirical research that the tail behaviour of certain financial series are time varying.

The tests considered by Phillips-Loretan and Koedjick *et al.* are based on Hill's conditional maximum likelihood estimator of the tail index of a distribution. Hill's estimator does not make use of the full sample to estimate the tail slope but only the m largest order statistics. This estimator has been shown to perform well in simulations (Kearns-Pagan, 1997). On theoretical grounds, it has been argued to be a more robust procedure for tail slope estimation compared with other estimators that use the full data and, in consequence, look at the centre of the distribution as well (DuMouchel, 1983). Furthermore, Hill's estimator and tests constructed from it do not require finiteness of the fourth moment.

Using Hill's estimator for tail thickness, we develop tests for the constancy of the tail behaviour over time when the breakdate is endogenous or *unknown*. While other papers have looked at the constancy of the tails using regression methods on the full sample, this is the first paper to construct tests for unknown breakpoints using an extreme value estimator. The framework of extreme value theory has the problem of selecting the m largest order statistics that goes into the estimation of the tails. In particular, the optimal selection rule for m has been shown to depend on properties of the tails (Hall, 1982). Because of this inherent circularity in tail slope estimation, DuMouchel (1983) suggested the simple rule that m be chosen as a fixed fraction of sample size. This rule has been shown to perform well in simulations and is widely used by practitioners in fields such as financial risk management, engineering, bankruptcy law and insurance.

It turns out that DuMouchel's rule can lead to the wrong test size for standard tests of structural change. When the fixed fraction is such that the corresponding m grows too fast with the sample size, then all three forms of the test considered here (recursive, rolling and sequential tests) diverge under the null. Under the alternative, the most commonly used test, the recursive test, which is based on the fluctuations test of Ploberger-Kramer-Kontrus (1989) fails as a two-sided test when m is optimal, but the rolling and sequential tests are consistent as two-sided tests. The power properties of these tests decrease with the wrong choice of m . This highlights the importance of finding an estimator for m rather than relying on a rule that disregards tail shape. We therefore also suggest a procedure for estimating m consistently.

We note that our test procedures are valid for both i.i.d. and serially dependent data. Since the current literature contains the theory of Hill's estimator for i.i.d. and linearly dependent data only, we derive in this paper the theory of Hill's estimator for nonlinear dependent data of the GARCH(1, 1) form. This theory draws heavily from Quintos (1999).

As an empirical application of our tests for time varying tail slope, we consider some recent data series whose tail behaviour is of much current interest and which have not, to our knowledge, been studied in the literature. Our sample of daily stock market returns covers the period from January 2, 1995 to October 16, 1998 which includes the recent Asian financial crisis. We apply our tests to stock market indices for the three Asian countries that were most affected by the crisis—Thailand, Malaysia and Indonesia. These three countries had their currency rates pegged to the U.S. Dollar and switched to a

floating rate system in July 1997. This known structural change in the operation of the foreign exchange market in these economies provides an interesting backdrop for the empirical application of our testing methods.

The rest of the paper is organized as follows. Section 2 presents an overview of the theory of Hill’s estimator for i.i.d. heavy tailed time series. Section 3 contains the tests for structural change and Section 4 extends the theory to accommodate serial dependence. Section 5 contains simulation results and Section 6 reports our empirical application to the emerging markets data. Section 7 offers some conclusions and the final section contains all proofs.

2. PRELIMINARIES

Consider the three time series of stock returns for Thailand, Malaysia and Indonesia shown in Figure 1. It is clear from these graphs that the degree of extreme movement in the series varies before and after July 1997, the time of the onset of the Asian financial crisis. One source of such variation in tail activity in a series is that the tail thickness of the underlying return distribution is not constant over time. Our aim is to develop test statistics that allow us to test the hypothesis of constant tail shape so we can address this question empirically in the case of the Asian financial crisis of 1997–1998.

To clarify ideas, we first consider an i.i.d. sequence $\{X_i\}_{i=1}^T$ whose tail behaviour is of the Pareto–Levy form,

$$\begin{aligned} P(X > x) &= acx^{-\alpha}(1 + dx^{-\beta} + o(x^{-\beta})), \\ P(X > -x) &= bcx^{-\alpha}(1 + dx^{-\beta} + o(x^{-\beta})), \end{aligned} \tag{1}$$

with $x > 0$ where $\alpha, c, \beta > 0$, $d \in \Re$ and the symmetry parameters, $a, b \geq 0$, satisfy $a + b = 1$. The parameter c is the scale dispersion parameter. The parameter that determines tail shape is the parameter α . It is also called the maximal moment exponent of the distribution since absolute moments of X of order less than α are finite.

When $0 < \alpha \leq 2$ then X is said to be “stable” Paretian in the sense that the maximal exponent does not change with convolution. The stable Paretian includes as a special case the Cauchy distribution ($\alpha = 1$) and the normal distribution ($\alpha = 2$). The case $0 < \alpha < 2$ is referred to as the “infinite variance” model. If $\alpha > 2$ then X is in the domain of attraction of a normal distribution and standardized partial sums of X converge in distribution to a normal.

We restrict our discussion to estimation and inference for the right tail of the distribution and set $a = 1$ (the behaviour of the left tail is easily inferred by premultiplication of a negative sign assuming symmetry). We let $X_{(i)}^J$ denote the i -th ordered statistic of the sample of size J , i.e. $X_{(1)}^J \leq \dots \leq X_{(J)}^J$. Hill’s estimator estimates α using m_J , the m largest observations of a sample of size J , $X_{(J-m_J+1)}^J \leq \dots \leq X_{(J)}^J$. For the full sample, Hill’s estimator is given by

$$\hat{\alpha}_T = \left(\frac{1}{m_T} \sum_{i=1}^{m_T} \log X_{(T-i+1)}^T - \log X_{(T-m_T+1)}^T \right)^{-1}. \tag{2}$$

The asymptotic properties of $\hat{\alpha}_T$ depend on the rate at which m_T grows with T . From Hall (1982) we have the following rates of convergence:

(A) If $m_T \rightarrow \infty$ and $m_T = o(T^{2\beta/(2\beta+\alpha)})$ then

$$m_T^{1/2}(\hat{\alpha}_T - \alpha) \xrightarrow{d} N(0, \alpha^2). \tag{3}$$

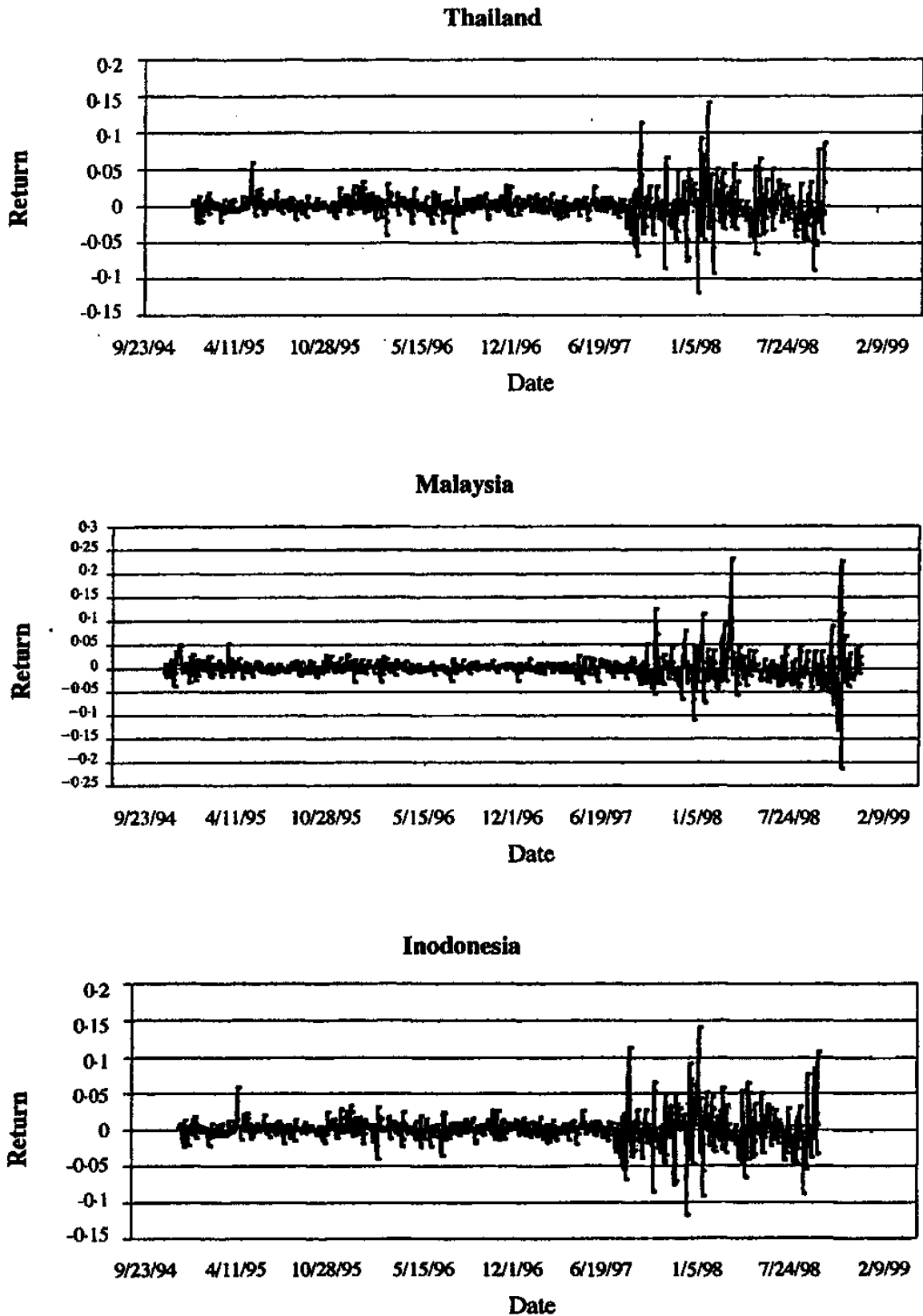


FIGURE 1

(B) If $m_T \sim \lambda T^{2\beta/(2\beta + \alpha)}$ for $0 < \lambda < \infty$ then

$$m_T^{1/2}(\hat{\alpha}_T - \alpha) \xrightarrow{d} N(\alpha d c^{-\beta/\alpha} \beta(\alpha + \beta)^{-1} \lambda^{(2\beta + \alpha)/2\alpha}, \alpha^2) = N(\alpha \phi, \alpha^2) \tag{4}$$

where $\lambda = \{c^{2\beta/\alpha}(1 + \beta/\alpha)^2/(2d^2(\beta/\alpha)^3)\}^{1/(2(\beta/\alpha) + 1)}$.

(C) If $d \neq 0$, $m_T = o(T)$ and $m_T/T^{2\beta/(2\beta + \alpha)} \rightarrow \infty$ then

$$(T/m_T)^{\beta/\alpha}(\hat{\alpha}_T - \alpha) \xrightarrow{p} d c^{-\beta/\alpha} \alpha \beta(\alpha + \beta)^{-1}.$$

As long as m_T grows slowly enough then Hill’s index converges to a normal distribution with variance α^2 . When m_T grows too fast then its distribution is degenerate.

Outcomes (A)–(C) show why tests for tail shape constancy have properties that depend on the choice of m_T . Note the circular problem in estimating α . We need m_T to estimate α in (2) but we also need α to select m_T in (A)–(C). DuMouchel (1983) suggested that m_T be chosen as a fraction of T , viz. $m_T = [\kappa T]$ with $\kappa = 0.10$, which implicitly selects an exponent for T . This presents a problem when the chosen κ leads to (B) or (C) and tests are constructed assuming (A) holds, as is usually done in practice.

We consider tests with a single unknown breakpoint and leave the extension to multiple breaks for later work. We let $t = [rT]$ for $r \in (0, 1)$ denote the *endpoint* of a subsample of size w_t . For example, for the full sample we have $w_t = w_T = T$. Correspondingly, we index Hill’s estimator with a time subscript as $\hat{\alpha}_t$ where

$$\hat{\alpha}_t = \left(\frac{1}{m_{w_t}} \sum_{i=1}^{m_{w_t}} \log X_{(w_t - i + 1)}^{w_t} - \log X_{(w_t - m_{w_t} + 1)}^{w_t} \right)^{-1}. \tag{5}$$

DuMouchel’s rule is then written as $m_{w_t} = [\kappa w_t]$.

The recursive, rolling and sequential estimators are modifications of (5). The recursive estimator $\hat{\alpha}_t$ is estimated from the subsample $[1, \dots, t]$. Thus, the recursive estimator is given by (5) with subsample size $w_t = t = [Tr]$. The rolling estimator fixes the subsample size w_t and estimates α using w_t rolled through time. Let $\gamma_0 \in (0, 1)$ denote the fraction of the fixed sample length and restrict $r \in (\gamma_0, 1)$. The calculation of the tail index α starts from $t_0 = [T(r - \gamma_0)] + 1$. Each subsample is of length $w_t^* = t - t_0 + 1 = [T\gamma_0]$ rolled through the full sample by eliminating past observations and adding future observations. Thus, the rolling estimator is (5) with subsample size w_t^* , viz.

$$\hat{\alpha}_t^* = \left(\frac{1}{m_{w_t^*}} \sum_{i=1}^{m_{w_t^*}} \log X_{(w_t^* - i + 1)}^{w_t^*} - \log X_{(w_t^* - m_{w_t^*} + 1)}^{w_t^*} \right)^{-1}.$$

The sequential test is constructed from a pre-break and a post-break estimator. The sequential pre-break estimator is just the recursive estimator. The post-break estimator, denoted by $\hat{\alpha}_{2t}$, is the reverse recursive estimator with sample size $w_t = w_{2t} = T - t$.

The following theorem gives invariance principles for the recursive, rolling and sequential estimators. The rates of convergence of (A) and (B) are also given with respect to the choice of κ .

Theorem 1. *Let $t = [Tr]$, $m_{w_t} = [\kappa w_t]$ and let $W(r)$ denote a standard Wiener process.*

(A) *If $m_{w_t} = o(w_t^{2\beta/(2\beta + \alpha)})$ or $\kappa = o(w_t^{2\beta/(2\beta + \alpha)^{-1}})$ then*

$$(i) \left(\frac{tm_t}{T} \right)^{1/2} (\hat{\alpha}_t - \alpha) \xrightarrow{d} \alpha W(r);$$

$$(ii) \left(\frac{w_t^* m_{w_t^*}}{T} \right)^{1/2} (\hat{\alpha}_t^* - \alpha) \xrightarrow{d} \alpha W(r, \gamma_0);$$

$$(iii) \left(\frac{w_{2t} m_{w_{2t}}}{T} \right)^{1/2} (\hat{\alpha}_{2t} - \alpha) \xrightarrow{d} \alpha W(1-r),$$

where $W(r, \gamma_0) = W(r) - W(s)$ with $s = r - \gamma_0$ in (ii).

(B)' If $m_{w_t} = \lambda w_t^{2\beta/(2\beta + \alpha)}$ or $\kappa = \lambda w_t^{2\beta/(2\beta + \alpha)^{-1}}$ then

$$(i) \left(\frac{tm_t}{T} \right)^{1/2} (\hat{\alpha}_t - \alpha) \xrightarrow{d} \alpha [W(r) + \varphi r^{1/2}];$$

$$(ii) \left(\frac{w_t^* m_{w_t^*}}{T} \right)^{1/2} (\hat{\alpha}_t^* - \alpha) \xrightarrow{d} \alpha [W(r, \gamma_0) + \varphi \gamma_0^{1/2}];$$

$$(iii) \left(\frac{w_{2t} m_{w_{2t}}}{T} \right)^{1/2} (\hat{\alpha}_{2t} - \alpha) \xrightarrow{d} \alpha [W(1-r) + \varphi(1-r)^{1/2}].$$

Conditions (A) and (B) can be derived as a special case of (i) with $t = T$ so $r = 1$. Condition (C)' requires that (C) be modified by replacing $T = w_t$ and $m_T = m_{w_t}$.

3. THE TESTS

Our null hypothesis is that the tail index α is constant over time. More specifically, let α_t be the tail index of the distribution of X_t . We focus on observations $t = [Tr]$ for $r \in R_\pi = [\pi, 1 - \pi]$, a prespecified compact subset of $(0, 1)$ for some small $\pi > 0$. The null hypothesis of constancy then takes the form

$$H_0: \alpha_{[Tr]} = \alpha \quad \forall r \in R_\pi \quad (6)$$

with the alternative $H_A: \alpha_{[Tr]} \neq \alpha$ for some $r \in R_\pi$. Sets like R_π are commonly used in the construction of parameter constancy tests (e.g. Hawkins, 1987, and Andrews, 1993) and represent some constant fraction of the overall number of observations, while being bounded away from zero and unity.

Since the optimal κ assumes knowledge of the tails in (B)' we assume that κ is set arbitrarily (say to 0.10) and assume that (A)' is satisfied. Our tests are constructed from the quantities

$$Y_T(t) = \left(\frac{tm_t}{T} \right)^{1/2} \left(\frac{\hat{\alpha}_t}{\hat{\alpha}_T} - 1 \right),$$

$$V_T(t) = \left(\frac{w_t^* m_{w_t^*}}{T} \right)^{1/2} \left(\frac{\hat{\alpha}_t^*}{\hat{\alpha}_T^*} - 1 \right),$$

$$Z_T(t) = \left(\frac{tm_t}{T} \right)^{1/2} \left(\frac{\hat{\alpha}_t}{\hat{\alpha}_{2t}} - 1 \right).$$

Note that the recursive test $Y_T(t)$ is based on the fluctuations test of Ploberger-Kramer-Kontrus since $((\hat{\alpha}_t/\hat{\alpha}_T) - 1) = \hat{\alpha}_T^{-1}(\hat{\alpha}_t - \hat{\alpha}_T)$. The same applies to the rolling test, since it

is the recursive test rolled through the full sample. The sequential test measures the fluctuation of the recursive estimator against the reverse recursive estimator as opposed to the full sample estimator.

Theorem 2. *Let $m_{w_t} = [\kappa w_t]$ and let (A)' hold. Then, under H_0 :*

- (i) $Q = \sup_{r \in R_n} Y_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_n} W(r)^2$
- (ii) $Q^* = \sup_{r \in R_n} V_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_n} W(r, \gamma_0)^2$
- (iii) $Q^\# = \sup_{r \in R_n} Z_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_n} W^\#(r)^2.$

where we have $\bar{W}(r) = W(r) - rW(1)$, $\bar{W}(r, \gamma_0) = W(r, \gamma_0) - (r - s)W(1, 1)$ and $W^\#(r) = W(r) - (r/(1 - r))W(1 - r)$.

The critical values of the tests are tabulated in Appendix A. With Theorem 1 and the CMT, it is possible to construct tests for structural change other than the sup test. Let $H_T(t)$ correspond to $Y_T(t)$, $V_T(t)$ or $Z_T(t)$ then the mean-score test and mean-exponential test of Andrews–Ploberger (1994) are, respectively,

$$\int_{R_n} H_T([Tr])^2 dr \xrightarrow{d} \int_{R_n} \Psi(r)^2 dr$$

and

$$\log \left\{ \int_{R_n} \exp(H_T([Tr])^2) dr \right\} \xrightarrow{d} \log \left\{ \int_{R_n} \exp(\Psi(r)^2) dr \right\}$$

where $\Psi(r)$ corresponds to the Wiener functionals given in Theorem 2.

Under the alternative hypothesis, we need the following result on the behaviour of Hill’s estimator with independent and non-identically distributed (i.n.i.d.) data. Let τ be the breakpoint and denote the subsamples as $\Upsilon_1 = [1, \tau]$ and $\Upsilon_2 = [\tau, T]$. The theorem below is for the case where the data is segmented into two parts (i.e. a single breakpoint), but the result will hold more generally for multiple breaks.

Theorem 3 (Hill’s Estimator under i.n.i.d. Data). *Let $\hat{\alpha}_T$ correspond to the estimate of α obtained from using the full sample $\{X_t\}$, $t \in [1, T] = \Upsilon$ (viz., equation (2)). For i.n.i.d. data $\{X_t\}$ such that $t \in \Upsilon_1$ has index α_1 and $t \in \Upsilon_2$ has index α_2 we have*

$$\hat{\alpha}_T \xrightarrow{d} \min(\alpha_1, \alpha_2) = \alpha. \tag{7}$$

According to (7), thick tails dominate in the sense that

$$\frac{P(X_{t \in \Upsilon_2} > x)}{P(X_{t \in \Upsilon_1} > x)} = (c_{\alpha_2}/c_{\alpha_1})x^{\alpha_1 - \alpha_2} \rightarrow 0,$$

as $x \rightarrow \infty$ iff $\alpha_1 < \alpha_2$ (and vice versa for $\alpha_2 < \alpha_1$). The implication of Theorem 3 is that Hill’s estimator converges to the dominant tail slope for samples w_t in which there are different tail indices.

The consistency of the tests in Theorem 2 now follow using Theorem 3.

Theorem 4. Assume (A)' is satisfied. Suppose τ is the breakpoint. We assume that $\alpha_t = \alpha_1$ for $t \in Y_1$ and $\alpha_t = \alpha_2$ for $t \in Y_2$. Then, under $H_A: H_{A1} \cup H_{A2}$,

$$\begin{aligned} H_{A1}: \alpha_1 > \alpha_2, & \quad H_{A2}: \alpha_1 < \alpha_2, \\ Q = O_p(m_t), & \quad Q = O_p(1), \\ Q^* = O_p(m_{w_t^*}), & \quad Q^* = O_p(m_{w_t^*}), \\ Q^* = O_p(m_t), & \quad Q^* = O_p(m_t). \end{aligned}$$

Case A1 has the tails varying from thinner to thicker and **Case A2** has the tails varying from thicker to thinner.

Remark 1. Under the alternative, the recursive test is consistent only under A1. It can therefore only be used as a one-sided test. The intuition behind this result is that Hill's estimator is conditional on the m largest observations, so that outlier behaviour that appears in the initial sample remains in the selection of the m observations in the latter part of the sample. This is not true for the rolling estimator since outlier behaviour that occurs in the initial sample is removed in the selection of the m largest observations as the sample is rolled through by eliminating past observations. The sequential test is consistent because it is constructed from the sum of the recursive and reverse recursive estimator, each of which is consistent in opposite directions.

Remark 2. Theorems 2 and 4 assume that κ is chosen so that (A)' is satisfied.

Corollary 5. Suppose κ satisfies (B)' then under H_0 :

$$\begin{aligned} \text{(i)} \quad Q &= \sup_{r \in R_\kappa} Y_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_\kappa} [W(r) + \varphi r^{1/2}(1-r^{1/2})]^2; \\ \text{(ii)} \quad Q^* &= \sup_{r \in R_\kappa} V_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_\kappa} [W(r, \gamma_0) + \varphi \gamma_0^{1/2}(1-\gamma_0^{1/2})]^2; \\ \text{(iii)} \quad Q^* &= \sup_{r \in R_\kappa} Z_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_\kappa} [W^*(r) + \varphi r^{1/2}(1-(r/(1-r))^{1/2})]^2. \end{aligned}$$

If κ satisfies (C)' then under H_0 ,

$$Q, Q^*, Q^* \rightarrow \infty.$$

If κ is such that m_{w_t} has the optimal rate, then all three tests have the wrong size if critical values under (A)' are used (*viz.* Appendix A). Furthermore, if κ is such that (C)' is satisfied then the tests diverge and the null is falsely rejected with probability one.

Remark 3. The current literature has tests of tail shape constancy that assume the sample splits are known. Suppose the sample size T is divided into g known subsamples according to a window length w_i , $\sum_{i=1}^g w_i = T$ and α is calculated for each subsample. We let subscripts $\{i\}$ denote the subsamples $i = 1, 2, \dots, g$. The null hypothesis with known breakpoint is

$$H_0: \alpha_{(1)} = \dots = \alpha_{(g)} = \alpha. \quad (8)$$

Koedijk *et al.* recommend the statistic

$$Q_0 = m_{(1)} \left(\frac{\hat{\alpha}_{(1)}}{\alpha_{(1)}} - 1 \right)^2 + \dots + m_{(g)} \left(\frac{\hat{\alpha}_{(g)}}{\alpha_{(g)}} - 1 \right)^2 \xrightarrow{d} \chi_g^2. \quad (9)$$

However, to implement this test one needs to specify α . Alternatively, one could use the test of Phillips–Loretan (1990) based on differences. For $g = 2$ and an equal sample split where $w_{(1)} = w_{(2)}$ they recommend the statistic

$$P_0 = \frac{m_{(1)}(\hat{\alpha}_{(1)} - \hat{\alpha}_{(2)})^2}{(\hat{\alpha}_{(1)}^2 + \hat{\alpha}_{(2)}^2)} \xrightarrow{d} \chi_1^2.$$

For an uneven sample split the extension for P_0 is straightforward. Assuming $m_{(1)} = [\gamma_{(2)}m_{(2)}]$ with $\gamma_{(2)} > 0$ we have

$$\begin{aligned} m_{(1)}^{1/2}(\hat{\alpha}_{(1)} - \hat{\alpha}_{(2)}) &= m_{(1)}^{1/2}(\hat{\alpha}_{(1)} - \alpha) - \left(\frac{m_{(1)}}{m_{(2)}}\right)^{1/2} m_{(2)}^{1/2}(\hat{\alpha}_{(2)} - \alpha) \xrightarrow{d} N(0, \alpha_{(1)}^2) - \gamma_{(2)}^{1/2} N(0, \alpha_{(2)}^2) \\ &= (\alpha_{(1)}^2 + \gamma_{(2)} \alpha_{(2)}^2)^{1/2} N(0, 1), \end{aligned} \tag{10}$$

so P_0 becomes

$$P_1 = \frac{m_{(1)}(\hat{\alpha}_{(1)} - \hat{\alpha}_{(2)})^2}{(\hat{\alpha}_{(1)}^2 + (m_{(1)}/m_{(2)})\hat{\alpha}_{(2)}^2)} = \frac{m_{(1)}\hat{\alpha}_{(2)}^2(\hat{\alpha}_{(1)}/\hat{\alpha}_{(2)} - 1)^2}{(\hat{\alpha}_{(1)}^2 + (m_{(1)}/m_{(2)})\hat{\alpha}_{(2)}^2)} \xrightarrow{d} \chi_1^2. \tag{11}$$

The revised test P_1 is the pointwise version of the sequential test $Q^*(r)$ when breakpoints are known and $Q^*(r)$ is normalized by its 2-period variance. The test Q_0 with $g = 1$ is the pointwise version of the recursive test with α estimated by $\hat{\alpha}_T$.

Remark 4. Value-at-Risk. Time-varying tails have implications for Value-at-Risk (VaR) calculation. VaR is defined as the maximum loss expected with $(1 - p)\%$ probability (typically taken as 5% or 1%) over a prespecified period of time (typically 1 day or 10 days). The VaR’s reported by financial institutions determine their capital requirements and also inform investors of the (extreme) risk associated with their investments. Under the extreme value theory framework, the effect of time-varying tails is clearly seen from the VaR formula,

$$\text{VaR}_p = \left(\frac{T}{m}(1 - p)\right)^{-1/\alpha} X_{(T-m+1)}.$$

Remark 5. Optimal Tests and Adaptation. The results in Theorems 2 and 4 are derived under rates that satisfy (A)’ and Remark 2 discusses the consequences of misspecification. It is possible to use the optimal rate (B)’ and to estimate m_w . Consider the following recursive estimator for m_t ,

$$\hat{m}_t = \hat{\lambda}_{t-1} w_t^{2\hat{\beta}_{t-1}/(2\hat{\beta}_{t-1} + \hat{\alpha}_{t-1})} = \hat{\lambda}_{t-1} w_t^{2\hat{\beta}_{t-1}/(2\hat{\beta}_{t-1} + 1)}, \quad \hat{\rho}_t = \hat{\beta}_t / \hat{\alpha}_t, \tag{12}$$

where $\hat{\beta}_{t-1} = \min(1, \hat{\alpha}_{t-1})$ and $\hat{\lambda}_{t-1} = (\hat{\epsilon}_{t-1}^{2\hat{\rho}_{t-1}}(1 + \hat{\rho}_{t-1})^2 / (2\hat{d}_{t-1}^2 \hat{\rho}_{t-1}^3))^{1/(2\hat{\rho}_{t-1} + 1)}$ with $\hat{\epsilon}_{t-1} = (\hat{m}_{t-1}/w_{t-1})(X_{(\hat{m}_{t-1}+1)}^*)^{\hat{\alpha}_{t-1}}$ and $\hat{d}_{t-1} = \hat{\epsilon}_{t-1}/2$. The initial selection of m_t can be done using a fixed fraction. Then, under the null hypothesis $\hat{m}_t - m_t \xrightarrow{p} 0$ where $m_t = \lambda w_t^{2\hat{\beta}/(2\hat{\beta} + \alpha)}$. Similarly, we can define the rolling estimator \hat{m}_{w_t} and the reverse recursive estimator \hat{m}_{w_t} .

Theorem 6. Assume (B)’ is satisfied and $\hat{m}_{w_t} - m_{w_t} \xrightarrow{p} 0$. Then

$$H = \sup_{r \in R_n} \left\{ H_T([Tr]) - \left(\frac{w_{[Tr]}}{T}\right)^{1/2} \left(\hat{\phi}_{w_{[Tr]}} - \left(\frac{\hat{m}_{w_{[Tr]}}}{\hat{m}_T}\right)^{1/2} \hat{\phi}_T \right) \right\}^2 \xrightarrow{d} \sup_{r \in R_n} \Psi(r)^2$$

where $H_T(t)$ is either $Y_T(t)$, $V_T(t)$ or $Z_T(t)$ and $\Psi(r)^2$ is the corresponding distribution in Theorem 2. For $Z_T(t)$ replace the subscript T by w_t , in the correction terms.

The correction term comes from the noncentral term φ in (B)', which is defined in (B). Note that with the choice of λ , φ simplifies to $\varphi = 1/\sqrt{2\beta/\alpha}$. Thus,

$$\hat{\phi}_{w_t} = \frac{1}{\sqrt{2\hat{\beta}_{w_t}/\hat{\alpha}_{w_t}}} \xrightarrow{P} \varphi_{w_t, R_n}$$

under H_0 where $\hat{\beta}_{w_t}$, $\hat{\alpha}_{w_t}$ correspond to the recursive, rolling or sequential estimates.

The time-varying estimators \hat{m}_t and \hat{m}_{w_t} "adapt" to the behaviour of the tail parameters α_t and β_t . They are alternatives to the adaptive estimator for m given by Hall–Welsh (1985) and to the bootstrap estimator of Danielsson–DeVries (1997a, b). The restriction $\beta_t = \min(1, \alpha_t)$ is the same as the restriction used by Hall–Welsh (1985) in adaptively estimating m . It is valid for stable distributions and symmetric distributions for which $\beta_t = \alpha_t$. An important distribution that the restriction excludes is the t -distribution where $\beta = 2$ and α_t equals the degrees of freedom.

4. TIME SERIES EXTENSIONS

The tests in Section 3 assume that $\{X_t\}$ is i.i.d. It is possible to modify the tests using Hsing's (1991) results for serially dependent data. Hill's estimator is still consistent and asymptotically normal with the same rates of convergence under serially dependent data, but with different variance. We let $x_+ = \max(x, 0)$ in what follows. Assume a consistent estimator for $X_{(T-m_T+1)}^T$ such that

$$\log X_{(T-m_T+1)}^T = \log b(T/m_T) + \delta/\sqrt{m_T}, \delta \in \mathfrak{R}$$

and $b(T/m_T) = F^{-1}(1 - m_T/T)$ is the quantile function which satisfies

$$\frac{T}{m_T} P(X > b(T/m_T)) \rightarrow 1. \quad (13)$$

Let

$$C_{T_i} = (\log X_i^T - \log b(T/m_T))_+, \quad \text{and} \quad D_{T_i} = I(\log X_i^T > \log b(T/m_T) + \delta/\sqrt{m_T}).$$

Lemma 7 (Hill's estimator for linearly dependent data; Hsing (1991), Theorem 3.3). *Let $\{X_t\}$ be a strictly stationary, l -dependent sequence with $l > 0$. Suppose there exist constants χ , ψ and ω such that for all $\delta \in \mathfrak{R}$,*

$$\begin{aligned} 2\alpha^2 \frac{T}{m_T} \sum_{j=2}^l EC_{T_i} C_{T_j} &\rightarrow \chi, \\ \alpha \frac{T}{m_T} \sum_{j=2}^l (EC_{T_i} D_{T_j} + EC_{T_j} D_{T_i}) &\rightarrow \psi, \\ 2 \frac{T}{m_T} \sum_{j=2}^l ED_{T_i} D_{T_j} &\rightarrow \omega. \end{aligned} \quad (14)$$

Then

$$m_T^{1/2}(\hat{\alpha}_T - \alpha) \xrightarrow{d} N(0, \alpha^2(1 + \chi + \omega - 2\psi)) = N(0, \alpha^2\eta), \quad (15)$$

under (A). Under (B), the same noncentral term enters into the limiting distribution.

We now extend Hsing’s result to the case where nonlinear dependence is allowed. Theorem 8 and Corollary 9 below draw heavily from Quintos (1999). We concentrate on the stationary ARCH and GARCH process which can be approximated by an l -dependent sequence. Let $\{X_t\}$ follow a GARCH(1, 1) process,

$$\begin{aligned} X_t &= \sigma_t Z_t, Z_t \sim \text{i.i.d. } N(0, 1), \\ \sigma_t^2 &= \beta_0 + \beta_1 \sigma_{t-1}^2 + \lambda X_{t-1}^2, \end{aligned} \tag{16}$$

with an ARCH(1) process being a special case with $\beta_1 = 0$. We work with a stochastic difference equation of the form

$$Y_t = A_t Y_{t-1} + B_t \tag{17}$$

where $\{A_t, B_t\}$ are i.i.d. pairs. By substitution, we can write (16) in terms of (17).

$$\sigma_t^2 = (\beta_1 + \lambda Z_{t-1}^2) \sigma_{t-1}^2 + \beta_0 = A_t \sigma_{t-1}^2 + B_t \tag{18}$$

with $Y_t = \sigma_t^2$, or for an ARCH(1),

$$X_t^2 = (\lambda Z_t^2) X_{t-1}^2 + \beta_0 Z_t^2 = A_t X_{t-1}^2 + \beta_t \tag{19}$$

with $Y_t = X_t^2$. Note that X_t^2 has tail index $\alpha/2$. Thus, when we construct Hill’s estimator in terms of Y_t we have

$$\hat{\alpha}_T = \left(\frac{1}{m_T} \sum_{i=1}^{m_T} \log Y_{(T-i+1)}^T - \log Y_{(T-m_T+1)}^T \right)^{-1} \xrightarrow{p} \alpha/2 = \hat{\alpha}.$$

We assume throughout that the conditions guaranteeing that (17) is strictly stationary, ergodic and regularly varying with tail index $\alpha/2$ are satisfied—viz. $\beta_0 > 0$ and $0 < \lambda < 1$ for an ARCH(1) and $\beta_0 > 0$ and $0 < \beta_1 + \lambda < 1$ for a GARCH(1, 1). We further assume that the process starts at Z_0 (i.e. $(Z_{-1}, \dots, Z_{-\infty}) = 0$). We can then write (17) in finite moving average form (with random coefficients),

$$Y_t = \prod_{i=1}^t A_i Y_0 + \sum_{j=1}^t \left(\prod_{i=j+1}^t A_i \right) B_j = G_1^t Y_0 + Y_t^{1,t} \tag{20}$$

where we have set $\prod_{i=j+1}^t A_i = 1$ for $j \geq t$. Note that stationarity requires $E(G_1^t) < 1$. We let $b = b(T/m_T)$ in what follows.

Theorem 8 (Hill’s Estimator for ARCH(1) and GARCH(1, 1) processes). *We construct Hill’s estimator in terms of the squared process $\{X_t^2\}$,*

$$\hat{\alpha}_T = \left(\frac{1}{m_T} \sum_{i=1}^{m_T} \log X_{(T-i+1)}^{2T} - \log X_{(T-m_T+1)}^{2T} \right)^{-1}.$$

Define

$$C_{T_i} = (\log X_i^{2T} - \log b)_+, \text{ and } D_{T_i} = I(\log X_i^{2T} > \log b + \delta/\sqrt{m_T}),$$

If $\{X_t\}$ follows an ARCH(1) (GARCH(1, 1)) process, then (a)–(c) ((d)–(f)) hold,

- (a) $2\hat{\alpha}^2(T/m_T) \sum_{j=2}^T EC_{T_j} C_{T_j} \rightarrow 2c \sum_{j=2}^{\infty} (2 + \hat{\alpha} E(G_1^{j-1})^\alpha) = \chi,$
- (b) $\hat{\alpha}(T/m_T) \sum_{j=2}^T (EC_{T_j} D_{T_j} + EC_{T_j} D_{T_j}) \rightarrow 2c \sum_{j=2}^{\infty} (1 + E(G_1^{j-1})^\alpha) = \psi,$
- (c) $2(T/m_T) \sum_{j=2}^T ED_{T_j} D_{T_j} \rightarrow 2c \sum_{j=2}^{\infty} (1 + E(G_1^{j-1})^\alpha) = \omega.$

- (d) $2\bar{\alpha}^2 T/m_T \sum_{j=2}^T EC_{T_1} C_{T_j} \rightarrow 2cE(Z_1^\alpha) \sum_{j=2}^\infty (2 + \bar{\alpha}E(G_1'^{-1})^\alpha) = \chi,$
 (e) $\bar{\alpha}(T/m_T) \sum_{j=2}^T (EC_{T_1} D_{T_j} + EC_{T_j} D_{T_1}) \rightarrow 2cE(Z_1^\alpha) \sum_{j=2}^\infty (1 + E(G_1'^{-1})^\alpha) = \psi,$
 (f) $2(T/m_T) \sum_{j=2}^T ED_{T_1} D_{T_j} \rightarrow 2c \sum_{j=2}^\infty E(Z_1^\alpha)(1 + E(G_1'^{-1})^\alpha) = \omega.$

From Lemma 7, Hill's estimator for an ARCH(1) or GARCH(1, 1) process is distributed as (15).

Lemma 7 and Theorem 8 show that Hill's estimator is asymptotically normal so that the tests in Section 3 need only be modified by scaling with a covariance matrix. Rather than construct empirical estimates of the right-hand terms in (a)–(f), we follow Hsing's method which uses empirical estimates for C_{T_i} and D_{T_j} . Define

$$c_{w_i} = (\log X_i^{2w_i} - \log X_{(w_i - m_{w_i} + 1)}^{2w_i})_+ \quad \text{and} \quad d_{w_i} = I(\log X_i^{2w_i} > \log X_{(w_i - m_{w_i} + 1)}^{2w_i})$$

and construct

$$\begin{aligned} \hat{\chi}_{w_i} &= (2\bar{\alpha}_{w_i}^2) \frac{1}{m_{w_i}} \sum_{j=1}^{w_i-1} c_{w_j} c_{w_{j+1}}, \\ \hat{\psi}_{w_i} &= (\bar{\alpha}_{w_i}) \frac{1}{m_{w_i}} \sum_{j=1}^{w_i-1} (c_{w_j} d_{w_{j+1}} + c_{w_{j+1}} d_{w_j}), \\ \hat{\omega}_{w_i} &= 2 \frac{1}{m_{w_i}} \sum_{j=1}^{w_i-1} d_{w_j} d_{w_{j+1}}, \end{aligned} \quad (21)$$

so that $\hat{\eta}_{w_i} = 1 + \hat{\chi}_{w_i} + \hat{\omega}_{w_i} - 2\hat{\psi}_{w_i}$. For example, for the full sample with $w_i = T$ we have

$$\begin{aligned} \hat{\chi}_T &= (2\bar{\alpha}_T^2) \frac{1}{m_T} \sum_{j=1}^{T-1} (\log X_j^{2T} - \log X_{(T-m_T+1)}^{2T})_+ (\log X_{j+1}^{2T} - \log X_{(T-m_T+1)}^{2T})_+, \\ \hat{\psi}_T &= (\bar{\alpha}_T) \frac{1}{m_T} \sum_{j=1}^{T-1} \{ (\log X_j^{2T} - \log X_{(T-m_T+1)}^{2T})_+ I(\log X_{j+1}^{2T} > \log X_{(T-m_T+1)}^{2T}) \\ &\quad + (\log X_{j+1}^{2T} - \log X_{(T-m_T+1)}^{2T})_+ I(\log X_j^{2T} > \log X_{(T-m_T+1)}^{2T}) \}, \\ \hat{\omega}_T &= 2 \frac{1}{m_T} \sum_{j=1}^{T-1} I(\log X_j^{2T} > \log X_{(T-m_T+1)}^{2T}) I(\log X_{j+1}^{2T} > \log X_{(T-m_T+1)}^{2T}). \end{aligned}$$

Note that χ , ψ and ω are zero if $\{X_i^2\}$ is i.i.d. Extensions of the results in the previous section now require normalization by $\hat{\eta}_{w_i}$.

Corollary 9. Suppose (A)' holds. Then, under H_0 :

- (i) $Q = \sup_{r \in R_n} \hat{\eta}_i^{-1} Y_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_n} W(r)^2;$
 (ii) $Q^* = \sup_{r \in R_n} \hat{\eta}_{w_i}^{-1} V_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_n} \bar{W}(r, \gamma_0)^2;$
 (iii) $Q^* = \sup_{r \in R_n} \hat{\eta}_{w_i}^{-1} Z_T([Tr])^2 \xrightarrow{d} \sup_{r \in R_n} W^*(r)^2,$

where $W(r)$, $\bar{W}(r, \gamma_0)$ and $W^*(r)$ are defined in Theorem 2. Under (B)', the distributional results of Corollary 5 holds.

Remark 6. The parameter α under the GARCH data generating process is a function of the parameters of the model. Thus, an alternative to Hill's estimator is to find α as the unique solution to

$$E(\lambda Z^2)^{\alpha/2} = 1 \quad \text{or} \quad E(\lambda Z^2 + \beta_1)^{\alpha/2} = 1 \tag{22}$$

for an ARCH(1) or GARCH(1, 1) respectively (see Embrechts, *et al.*, 1997, and Mikosch-Starica, 1998). For example, the relation of the tail index of an ARCH(1) process and its conditional parameter λ is tabulated in Embrechts *et al.* (1997) and is repeated here.

λ	0.3125	0.4	0.5	0.5773	0.6	0.7	0.8	0.9	1	1.57
α	8.00	6.09	4.74	4.00	3.82	3.17	2.68	2.30	2.00	1.00

We observe, in particular, that the second moment is infinite when $\lambda \geq 1$.

Remark 7. (Conditional versus Unconditional Tail Testing). An important implication of (22) is that instability in the ARCH/GARCH parameter λ implies instability in the tail parameter α of the unconditional distribution, and vice versa. Thus, the *unconditional* distribution of $\{X_t\}$ has very fat tails when λ is greater than 1. Our *conditional* distribution has Gaussian (thin) tails irrespective of λ , *viz.*

$$f(X_t | \sigma_t) \equiv Z_t \sim N(0, 1).$$

The relation of λ and α given in Remark 6 relies on this assumption. Our limit theory is not justified by the analysis given here when thick tails appear in the conditional distribution, for example, when $\{Z_t\}$ is *t*-distributed. Such cases may be analysed by an extension of our methods but we have not done so in this paper.

5. SIMULATIONS

This section explores the finite sample properties of the parameter constancy tests using 5000 Monte Carlo simulations. Sections 5.1 and 5.2 consider only the i.i.d. case to highlight the results of Theorem 4. Section 5.3 extends the simulations to the dependent case. All simulations were done using Visual C++ version 4.0.

5.1. Data generating process

We split our sample into two periods of length $T_1 = [rT]$ ($r \in (0, 1)$) and $T_2 = T - T_1 + 1$. For each sub-period, we generate the X_t 's randomly from a stable distribution with index α using the method of Samorodnitsky–Taqqu (1994), *viz.* we set

$$X_t = \frac{\sin \alpha \gamma_t}{(\cos \gamma_t)^{1/\alpha}} \left(\frac{\cos((1 - \alpha)\gamma_t)}{W_t} \right)^{(1-\alpha)/\alpha},$$

where γ is uniform on $(-\pi/2, \pi/2)$ and W is exponential with mean 1. Different values of the shape parameter α were chosen to lie in the interval $0 < \alpha < 1$ (where the mean and variance are infinite) and $1 < \alpha < 2$ (where the variance is infinite) for each period. Both Cases A1 and A2 are tested, *i.e.* from thin to thick tails and from thick to thin tails. Breakpoints are imposed *ex ante*. Specifically, we set $r = 0.25, 0.5$ and 0.75 . Sample size T is set to 500, and 2000, two sample sizes representative of those commonly found in financial data applications. We use 5000 iterations.

TABLE 1
Size of $Q(r)$, $Q^*(r)$, $Q^{\dagger}(r)$

	$T = 500$		$T = 2000$	
	Nominal size		Nominal size	
	0.01	0.05	0.01	0.05
Recursive				
$\alpha = 0.5$	0.02	0.08	0.01	0.06
$\alpha = 1.5$	0.01	0.06	0.01	0.05
Rolling ($\gamma_0 = 0.15$)				
$\alpha = 0.05$	0.03	0.07	0.02	0.07
$\alpha = 1.5$	0.02	0.07	0.01	0.06
Sequential				
$\alpha = 0.5$	0.03	0.08	0.01	0.07
$\alpha = 1.5$	0.02	0.06	0.01	0.05

5.2. Structural change tests for α

Since we specified α in our simulation we can ensure that (A)' is satisfied. The problem then is how to select m since both estimators $\hat{\alpha}$ and \hat{c} are constructed conditional on the m largest observations. DuMouchel (1983) suggests the simple rule $m \leq 0.1T$, i.e. that the upper 10 percentile be used as extreme values. Hall (1982) and Hall-Welsh (1985) find the optimal rate to be $m = 2T^{2/3}$. In this paper, we set $\kappa = 0.10$ and select α such that $m_{w_t} = [0.10w_t] < \lambda_t w_t^{2\beta_t/(2\beta_t + \alpha_t)}$.

Table 1 gives the size properties of the tests. Note the good size properties in both the finite and infinite variance case. The tests perform better the further out in the tails, as expected from extreme value theory.

Table 2(a) reports the power properties of the recursive test. For the power property of the tests, we look at 2 cases: Case A1—where the change goes from thin to thicker tails and Case A2—where the tails become thinner as T gets larger. Notice the drop in power for Q from Case A1 to Case A2. This is, in fact, intuitively reasonable since the tests are based on estimators that use only the *largest* observations. In Case A2, where the thicker tails occur in the initial period, the outliers in this tail enter into the calculation of successive tests as long as they belong in the m largest observations. For example, if the maximum point in the data occurs in the first subsample, this point will enter into the calculation of each successive test, thereby making the tails appear thicker than they really are. This will not be the case when the outlier behaviour occurs in the latter part of the sample, as evidenced by Case A1. These results support the theoretical findings in Theorem 4.

TABLE 2A
Power properties of $Q(r)$ (nominal size = 0.05)

	(α_1, α_2) Power	$T = 500$			$T = 2000$		
		Breakpoints			Breakpoints		
		$r = 0.25$	$r = 0.50$	$r = 0.75$	$r = 0.25$	$r = 0.50$	$r = 0.75$
A1	(0.8, 0.5)	0.3436	0.4798	0.3676	0.5426	0.8724	0.7450
	(1.8, 1.2)	0.6112	0.7600	0.6050	0.9362	0.9968	0.9606
	(1.5, 0.8)	0.6864	0.8790	0.7930	0.9852	0.9998	0.9988
A2	(0.5, 0.8)	0.0078	0.0172	0.0302	0.0000	0.0002	0.0008
	(1.2, 1.8)	0.0020	0.0104	0.0268	0.0028	0.0112	0.0008
	(0.8, 1.5)	0.0042	0.0130	0.0340	0.0048	0.0002	0.0004

TABLE 2B
Power properties of $Q^*(r)$ (nominal size = 0.05)

		T = 500			T = 2000		
		Breakpoints			Breakpoints		
(α_1, α_2)	Power	r = 0.25	r = 0.50	r = 0.75	r = 0.25	r = 0.50	r = 0.75
A1	(0.8, 0.5)	0.4792	0.5616	0.5248	0.3020	0.3722	0.1740
	(1.8, 1.2)	0.6372	0.7216	0.6234	0.7582	0.7952	0.3958
	(1.5, 0.8)	0.7114	0.8362	0.7642	0.9034	0.9656	0.7500
A2	(0.5, 0.8)	0.5198	0.5724	0.4760	0.1752	0.3650	0.3200
	(1.2, 1.8)	0.6300	0.7206	0.6348	0.4014	0.7968	0.7652
	(0.8, 1.5)	0.7784	0.8328	0.7024	0.7524	0.9608	0.9092

Tables 2(b) and (c) report the power properties of the rolling test Q^* and sequential test Q^s . Compared with the recursive test, the rolling test has more power in both directions: from thick to thin tails, and from thin to thick tails which coincides with Theorem 4. The sequential test seems to do poorly when the sample size is small ($T = 500$). Its power also depends on the location of the breakpoint and the direction that the tail varies. This is, of course, due to the fact that the test is constructed from two statistics which are consistent in opposite directions. At $T = 2000$ for A1, the recursive test is consistent (and the reverse recursive test is inconsistent) but the power of the test is low because of sample size (*i.e.* only 25% of the sample is used for the recursive test). At 0.75 the power of the recursive test dominates. The symmetry in the power of A1 and A2 depending on the breakpoint is evident in the tables.

Table 3 reports estimates of the breakpoint, *i.e.* the location of the sup value of the test. The recursive test does well in estimating the breakpoint for Case A1. For Case A2, the sup values occur far from the location of the true breakpoints. Since the test is inconsistent in this direction the location of the breakpoint is also inconsistent. The breakpoint estimates of Q improve as T becomes larger, both in terms of mean and standard error.

While the rolling test is consistent, its estimate of the breakpoints do not coincide with the true breakpoints. The sequential test Q^s , however, seems to perform well in estimating the breakpoints for both Case A1 and A2 when $T = 2000$. Its weakness lies in its poor small sample behaviour. In our example, it fails when $T = 500$.

5.3. Extension to ARCH(1)

Our data generating process is an ARCH(1) as given in (16). We set $\beta_0 = 1$ and vary λ since, from Remark 6, this gives the value for the tail index α . Table 4 contains the

TABLE 2C
Power properties of $Q^*(r)$ (nominal size = 0.05)

		T = 500			T = 2000		
		Breakpoints			Breakpoints		
(α_1, α_2)	Power	r = 0.25	r = 0.50	r = 0.75	r = 0.25	r = 0.50	r = 0.75
A1	(0.8, 0.5)	0.0002	0.0010	0.0086	0.0110	0.2082	0.3074
	(1.8, 1.2)	0.0012	0.0078	0.0070	0.2372	0.8496	0.8562
	(1.5, 0.8)	0.0002	0.0060	0.0874	0.3388	0.9376	0.9494
A2	(0.5, 0.8)	0.0116	0.0020	0.0000	0.3006	0.2056	0.0116
	(1.2, 1.8)	0.0676	0.0074	0.0008	0.8528	0.8452	0.2348
	(0.8, 1.5)	0.0866	0.0054	0.0002	0.9518	0.9372	0.3440

TABLE 3

Estimates of breakpoints

(α_1, α_2) Power	$T = 500$			$T = 2000$		
	Breakpoints			Breakpoints		
	$r = 0.25$	$r = 0.50$	$r = 0.75$	$r = 0.25$	$r = 0.50$	$r = 0.75$
Recursive test						
A1 (0.8, 0.5)	0.2897 (0.1195)	0.4492 (0.1446)	0.5908 (0.2140)	0.2687 (0.0650)	0.4945 (0.0663)	0.7031 (0.0984)
(1.8, 1.2)	0.2826 (0.1075)	0.4646 (0.1296)	0.6162 (0.1978)	0.2608 (0.0477)	0.4971 (0.0468)	0.7182 (0.0746)
(1.5, 0.8)	0.2653 (0.0848)	0.4654 (0.1166)	0.6500 (0.1753)	0.2509 (0.0284)	0.4953 (0.0335)	0.7306 (0.0510)
A2 (0.5, 0.8)	0.2922 (0.1956)	0.2268 (0.0961)	0.2540 (0.1256)	—	0.1625 (—)	0.3388 (0.1766)
(1.2, 1.8)	0.3084 (0.2578)	0.2242 (0.1160)	0.2442 (0.1130)	0.4898 (0.0551)	0.5276 (0.0384)	0.1878 (0.0103)
(0.8, 1.5)	0.1941 (0.0520)	0.2126 (0.0676)	0.2668 (0.1236)	0.4709 (0.0519)	0.5550 (—)	0.2243 (0.0712)
Rolling test						
A1 (0.8, 0.5)	0.3868 (0.2509)	0.3966 (0.1907)	0.4821 (0.2028)	0.2129 (0.0661)	0.3336 (0.1145)	0.4566 (0.1780)
(1.8, 1.2)	0.3161 (0.2147)	0.3633 (0.1574)	0.4676 (0.1966)	0.2070 (0.0443)	0.3293 (0.1151)	0.4543 (0.1828)
(1.5, 0.8)	0.2941 (0.1978)	0.3589 (0.1538)	0.4620 (0.1924)	0.2060 (0.0412)	0.3305 (0.1138)	0.4537 (0.1843)
A2 (0.5, 0.8)	0.6619 (0.2024)	0.7387 (0.1918)	0.7494 (0.2530)	0.6897 (0.1823)	0.8195 (0.1185)	0.9366 (0.0710)
(1.2, 1.8)	0.6774 (0.1949)	0.7828 (0.1591)	0.8386 (0.2058)	0.6909 (0.1833)	0.8157 (0.0132)	0.9444 (0.0418)
(0.8, 1.5)	0.6781 (0.1934)	0.7874 (0.1508)	0.8511 (0.1975)	0.6951 (0.1844)	0.8196 (0.1152)	0.9460 (0.0417)
Sequential test						
A1 (0.8, 0.5)	0.8240 (—)	0.8086 (0.0367)	0.8113 (0.0393)	0.3469 (0.1803)	0.5254 (0.0827)	0.7416 (0.0656)
(1.8, 1.2)	0.8310 (0.0246)	0.7735 (0.0806)	0.7978 (0.0434)	0.2906 (0.0919)	0.5182 (0.0594)	0.7505 (0.0501)
(1.5, 0.8)	0.8440 (—)	0.7093 (0.1037)	0.7886 (0.0459)	0.2599 (0.0343)	0.4987 (0.0391)	0.7264 (0.0557)
A2 (0.5, 0.8)	0.1926 (0.0411)	0.2042 (0.0491)	—	0.2586 (0.0650)	0.4781 (0.0799)	0.6843 (0.1458)
(1.2, 1.8)	0.1999 (0.0416)	0.2074 (0.0679)	0.1920 (0.0279)	0.2497 (0.0499)	0.4836 (0.0565)	0.7122 (0.0860)
(0.8, 1.5)	0.2145 (0.0438)	0.2506 (0.1064)	0.1740 (—)	0.2737 (0.0563)	0.5018 (0.0386)	0.7397 (0.0389)

1. All computations of Hill index use $k = 0.10$.2. Table shows mean estimate of r over 5000 iterations. Standard errors are in parentheses.

TABLE 4

Size and power properties of tests of ARCH(1) process

$(\alpha_1, \alpha_2) =$	Size Nominal size = 0.05			
	$\alpha = 1$		$\alpha = 4$	
	Case A1	Case A2	Case A1	Case A2
Recursive	0.04	0.01	0.04	0.01
Rolling	0.05	0.03	0.05	0.03
Sequential	0.04	0.01	0.04	0.01
Power				
	Case A1	Case A2	Case A1	Case A2
$r = 0.25$	0.8970	0.4260	0.1886	0.0500
$r = 0.5$	0.9400	0.5280	0.2662	0.0938
Rolling				
$r = 0.25$	0.4500	0.2400	0.4300	0.2600
$r = 0.5$	0.5500	0.3621	0.5484	0.3560
Sequential				
$r = 0.25$	0.0001	0.0000	0.0002	0.0000
$r = 0.5$	0.0020	0.0000	0.0030	0.0001

*Notes:** Sample size $T = 500$.

** For rolling test, rolling sample size is 20% of full sample size.

simulation outcomes for the size and power properties of our tests. We give results only for $T = 500$ since a larger sample will give only better results. The tests have the correct size but the power is lower than in Table 2 for the i.i.d. case. The pattern of inconsistency that appears in Table 2 can be seen here. The recursive test has significantly lower power for Case A2 than Case A1 and this is revealed as well in the sequential test. Recall that the sequential test is constructed from the recursive and reverse recursive test. The rolling test has the best power in both directions as expected. The sample size fraction used in these simulations is $\gamma_0 = 0.20$.

6. APPLICATION TO THE TAIL BEHAVIOUR OF ASIAN FINANCIAL MARKETS

6.1. Data

Our dataset consists of the daily stock price index series of Thailand, Malaysia and Indonesia, downloaded from Bloomberg Financial Markets. The sample covers the period from January 2, 1995 to October 16, 1998, which includes 931 daily return observations for the Bangkok SET Index of Thailand (SET), 933 daily return observations for the Kuala Lumpur Composite Index of Malaysia (KLCI) and 937 daily return observations for the Jakarta Composite Index (JCI). This sample covers the period of the recent financial crisis.

These stock market data present interesting cases for studying tail behaviour as all three countries experienced financial economic turmoil since May 1997. Following IMF recommendations, all three countries switched from a pegged exchange rate system to a floating rate system in July 1997. Our interest is in whether equity markets reacted differently under the two currency regimes in terms of extreme movements in equity returns. When exchange rates float, external shocks are partly absorbed by currency exchange rate fluctuations. The institution of a floating currency should therefore play a role in influencing the extent of extreme price movements in other markets, such as the domestic equity market.

6.2. Results

We are interested in whether there was a change in the tail behaviour of the stock market during this period. A rise (fall) in α would imply a reduction (an increase) of extreme market movements. Figure 1 graphs changes in the stock prices and seems to indicate that all three countries exhibit a change in the degree of extreme movements over time. We first apply our version of the Phillips–Loretan test of unequal sample split (*viz.* P_1) on the three data samples using July 1997 as our breakpoint to test this assertion, then perform the sequential, recursive and rolling tests assuming no knowledge of the breakpoints. We use our most general result, Theorem 8 and Corollary 9, throughout this exercise (Theorems 8 and 9 reduce to the i.i.d. tests of Section 3). The test size is set to 5%.

6.2.1. Sequential test

We split each data series into two periods according to the dates when each of the three countries announced that they would no longer support their pegged exchange rate systems. All countries widened the trading band first before they let their currencies float. We treat the widening of the trading band as an interim step toward the float system.

Therefore, we picked the initial central bank announcement date of widening the trading band as the breakpoint for our sample. Specifically, the SET index is divided into two subsamples by July 2, 1997 when the Bank of Thailand announced a managed float of the baht. KLCI index is separated on July 14, 1997 when the Malaysian central bank, Bank Negara, abandoned the defence of the ringgit, and we split the JCI index into two by July 11, 1997 when Jakarta widened its rupiah trading band from 8% to 12%.

TABLE 5

The tail index of three equity markets in two sub-periods

Equity index	Pegged rate system Point estimate	Floating rate system Point estimate	P_1
SET	2.54 (2.10–2.99)	2.15 (1.51–2.78)	1.51
KLCI	2.51 (2.03–3.00)	1.84 (1.32–2.43)	5.12
JCI	2.80 (2.26–3.33)	2.05 (1.51–2.60)	5.37

The results for a known breakpoint are reported in Table 5. The first and second column give Hill's estimates of the tail index for the pre- and post-break periods. The numbers in parentheses are the corresponding 95% confidence interval. From Theorem 8, our confidence interval is constructed as

$$\hat{\alpha}_{T_i} \pm (z_{\text{size}/2}) \left[\frac{\hat{\alpha}_{T_i} \eta_{T_i}}{m_{T_i}} \right]^{1/2}, \quad i = 1, 2$$

where $z_{\text{size}/2} = z_{0.05/2} = 1.96$, i denotes the pre and post sample periods and T_i the corresponding sample size. Note that all three indices are higher in the first period and the confidence intervals show that all three indices have a point estimate of α above 2 under this regime. The estimates of the second period are lower and are indicative of fatter tails.

The revised Phillips–Loretan test P_1 for the null hypothesis that $\alpha_1 = \alpha_2$ for all three series is given in the third column. Following the notation of Section 3, the test is revised to accommodate dependence as

$$Q_1 = \frac{m_{(1)} (\hat{\alpha}_{(1)} - \hat{\alpha}_{(2)})^2}{(\hat{\alpha}_{(1)}^2 \eta_{(1)} + (m_{(1)}/m_{(2)}) \hat{\alpha}_{(2)}^2 \eta_{(2)} - 2\eta_{(12)})} \xrightarrow{d} \chi_1^2. \quad (23)$$

The variances $\eta_{(1)}$ and $\eta_{(2)}$ are calculated using (21) with corresponding sample size T_1 and T_2 . The covariance term $\eta_{(12)}$ is constructed from

$$\hat{\chi}_{T_n} = \frac{2\hat{\alpha}_{\min}^2}{m_{\min}} \sum_{j=1}^{T_n^*} \sum_{i=1}^{T_x} (\log X_j^{2T_n^*} - \log X_{(T_n^* - m_{T_n^*} + 1)}^{2T_n^*})_+ (\log X_i^{2T_x} - \log X_{(T_x - m_{T_x} + 1)}^{2T_x})_+,$$

$$\hat{\psi}_{T_n} = \frac{\hat{\alpha}_{\min}}{m_{\min}} \sum_{j=1}^{T_n^*} \sum_{i=1}^{T_x} (\log X_j^{2T_n^*} - \log X_{(T_n^* - m_{T_n^*} + 1)}^{2T_n^*})_+ I(\log X_i^{2T_x} > \log X_{(T_x - m_{T_x} + 1)}^{2T_x}),$$

$$\hat{\omega}_{T_n} = \frac{2}{m_{\min}} \sum_{j=1}^{T_n^*} \sum_{i=1}^{T_x} I(\log X_j^{2T_n^*} > \log X_{(T_n^* - m_{T_n^*} + 1)}^{2T_n^*}) I(\log X_i^{2T_x} > \log X_{(T_x - m_{T_x} + 1)}^{2T_x}),$$

where $T_x = \max(T_1, T_2)$, $T_n = \min(T_1, T_2)$, $\hat{\alpha}_{\min} = \min(\hat{\alpha}_{(1)}, \hat{\alpha}_{(2)})$ and m_{\min} is the number of order statistics used to calculate $\hat{\alpha}_{\min}$. With a χ_1^2 critical value of 3.84, the test P_1 rejects the null of tail constancy for both Malaysia and Indonesia at the 5% level.

Figure 2 plots the sequential test Q^* for all three series. From Appendix A, the 95% critical value is 18.31 and the 90% critical value is 13.98. Figure 2 shows that we are once again able to reject the null of tail constancy for Malaysia, but not for Indonesia even at

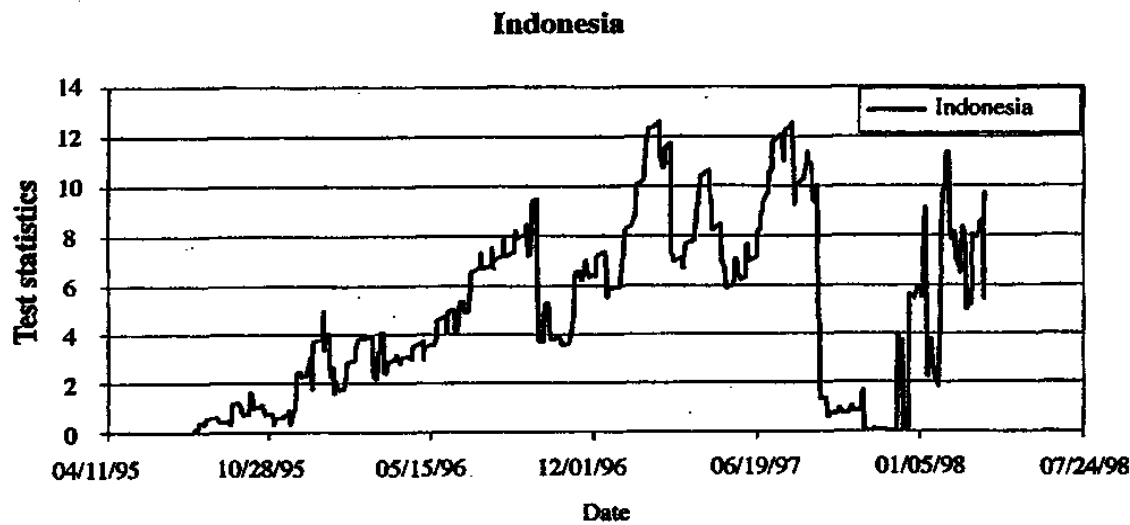
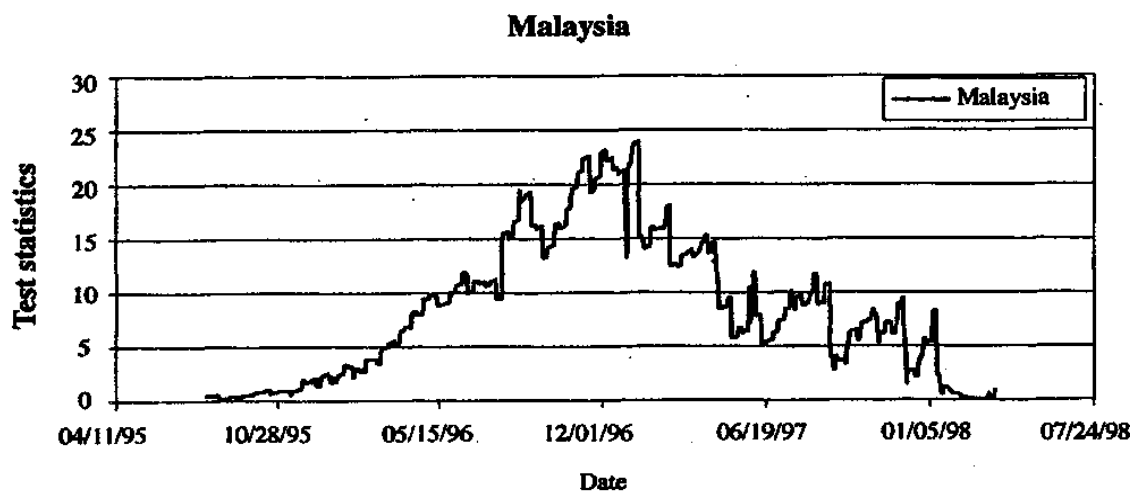
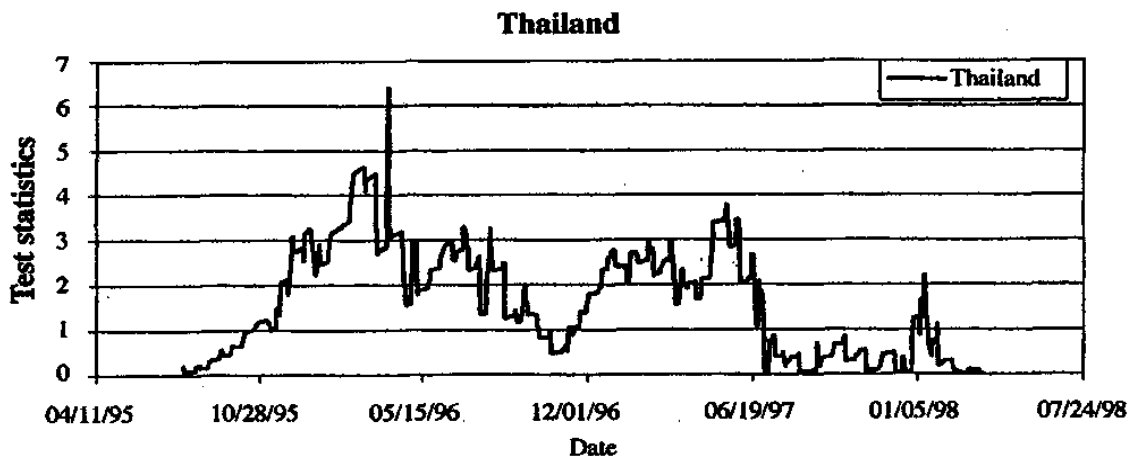


FIGURE 2
Sequential test (95% critical value = 18.31)

the 10% level (test statistic is 12.60). A possible reason for the failure of the test is that we do not have a large sample size and we know from the simulations that Q^* does not perform well in small samples. The estimated breakpoints are listed in Table 5.

6.2.2. Recursive test

For our recursive test of Theorem 2, we start with a window size of 500 observations and successively increase the sample size by a single observation. Figure 3 shows the evolution of the tail index starting from January 1997. It shows that the tail index for all three series trend downward through the period and stays relatively stable at a lower level during 1998. The transition takes place in the mid and late half of 1997, which is roughly about the time when the Asian financial crisis occurred. The downward trend of α indicates that there is an increase in extreme movement in all three markets during the period.

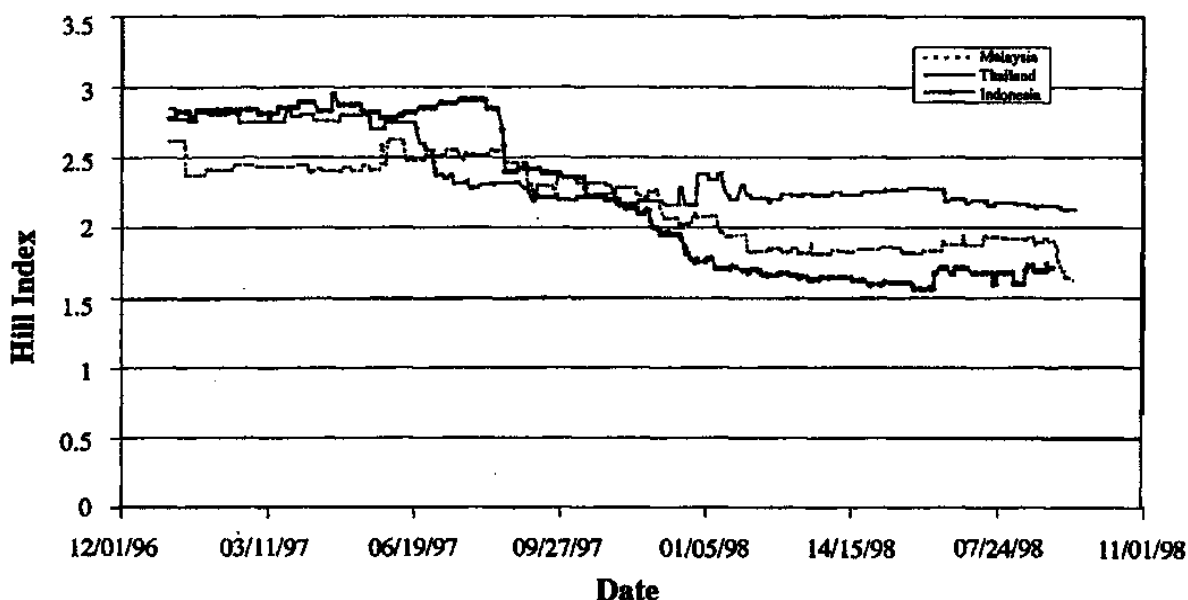
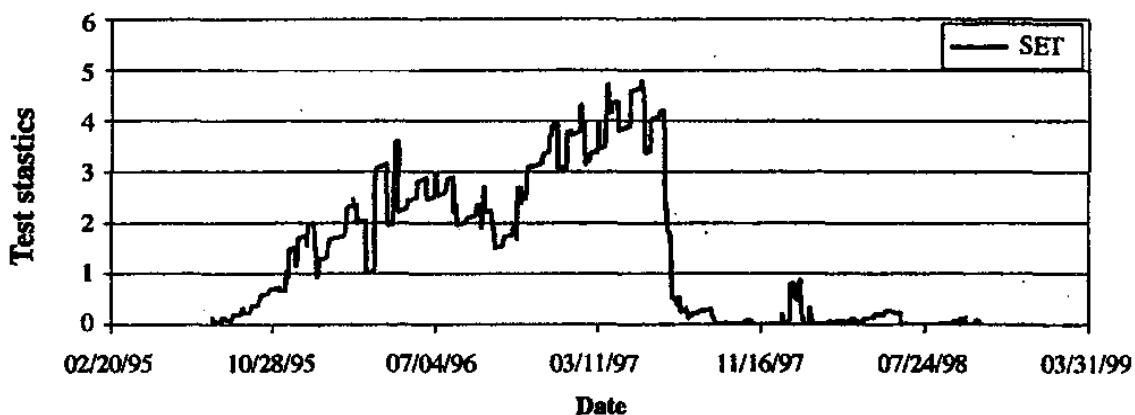


FIGURE 3
Hill's index: recursive

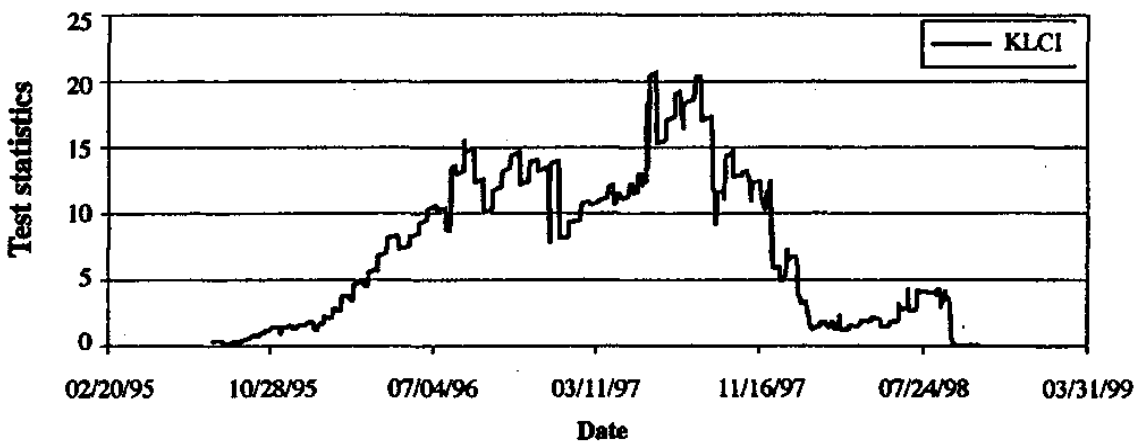
The recursive test statistic Q for all three equity indices during the period are plotted in Figure 4. We are able to strongly reject the one sided null of constant α (i.e. $H_0: \alpha_t \leq \alpha_T$ vs. $H_A: \alpha_t > \alpha_T$) for all three indices, with much stronger rejection for both Malaysia and Indonesia. Suprema of the test statistics and hence the breakpoint dates detected by the test are reported in Table 6. Note that for Malaysia and Indonesia, the selected breakpoints are close to July 1997.

6.2.3. Rolling test

The rolling test Q^* is a modification of the recursive test, for which we keep the window size fixed at 500 observations and roll it through the sample. Our results are summarized in Figures 5 and 6. Figure 5 confirms the downward trend of α . The difference between the rolling α and the recursive α is that the extreme movements during the previous period will be carried over in calculating the recursive α , but not so much in calculating the rolling α . The chart suggests that there tends to be an increase in extreme movements in all three equity markets.



Malaysia



Indonesia

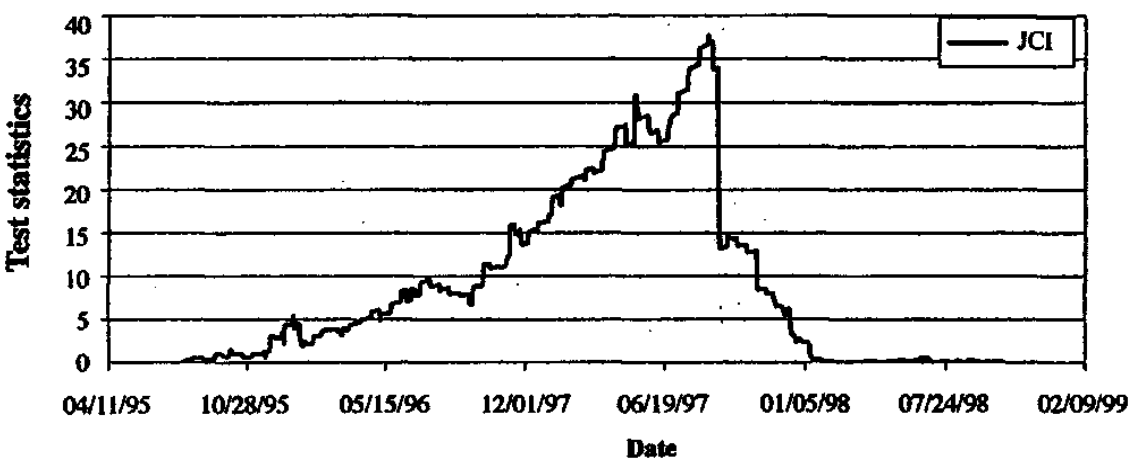


FIGURE 4
Recursive test (95% critical value = 1.78)

TABLE 6

Tests of breakpoints

Equity index	Recursive test		Rolling test		Sequential test	
	Sup	Date	Sup	Date	Sup	Date
SET	4.78	16/05/97	7.49	03/06/97	6.39	04/04/96
KLCI	20.37	21/08/97	53.75	19/08/97	23.90	13/01/97
JCI	37.74	20/08/97	37.57	14/08/97	12.60	19/02/97

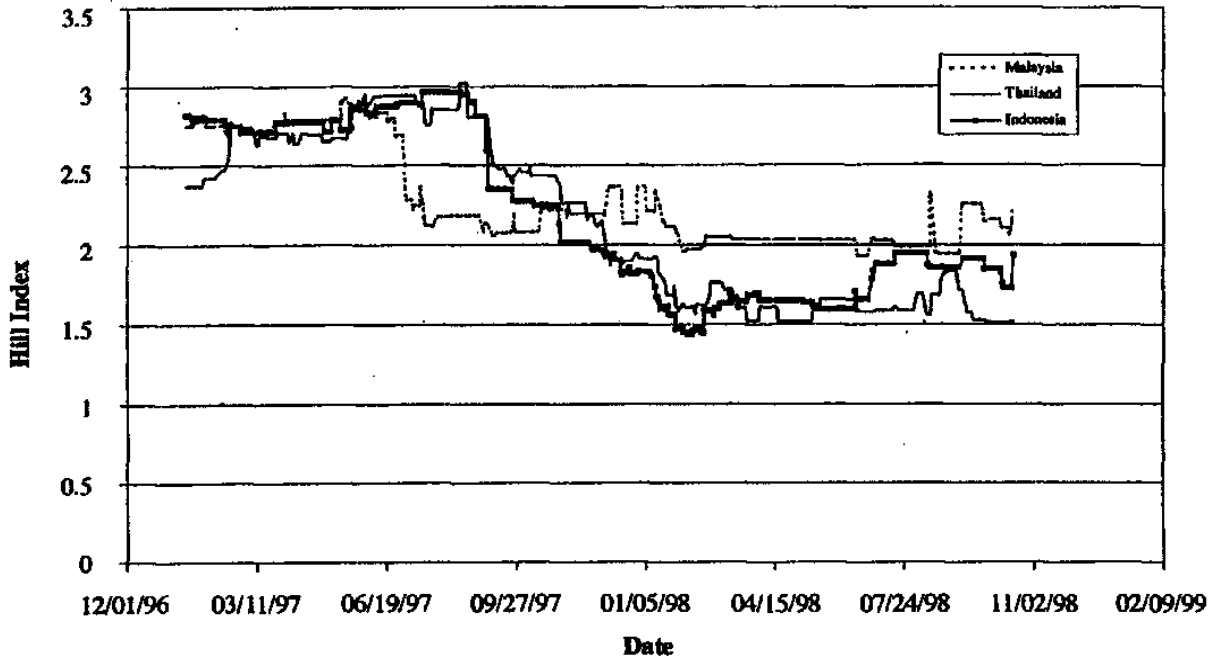


FIGURE 5
Hill's index: rolling

The rolling test statistic Q^* for all three data series are plotted in Figure 6. Suprema of the test statistic and the breakpoint dates detected by the test are again summarized in Table 6. Like the recursive test, the rolling test strongly rejects the null hypothesis of constant α for both KLCI and JCI indices. Moreover, the breakpoints picked by the rolling test are close to those picked by the recursive test. Note that the rolling test produces more "flat" surfaces due to the construction of the test so that it can reach its suprema on multiple dates.

7. CONCLUSION

This paper has constructed tests for tail shape constancy using Hill's estimator. It is the first to do so with unknown breakpoints. The results show that standard tests for structural change have quite different power properties in this context. Furthermore, unless the tests are normalized at the correct rate, the tests can diverge under the null and have no power. This highlights the importance of the optimal/adaptive tests discussed in Remark 5.

The simulations corroborate the limit theory that the recursive test is a one-sided test. For cases where it is valid, the recursive test outperforms the rolling and sequential

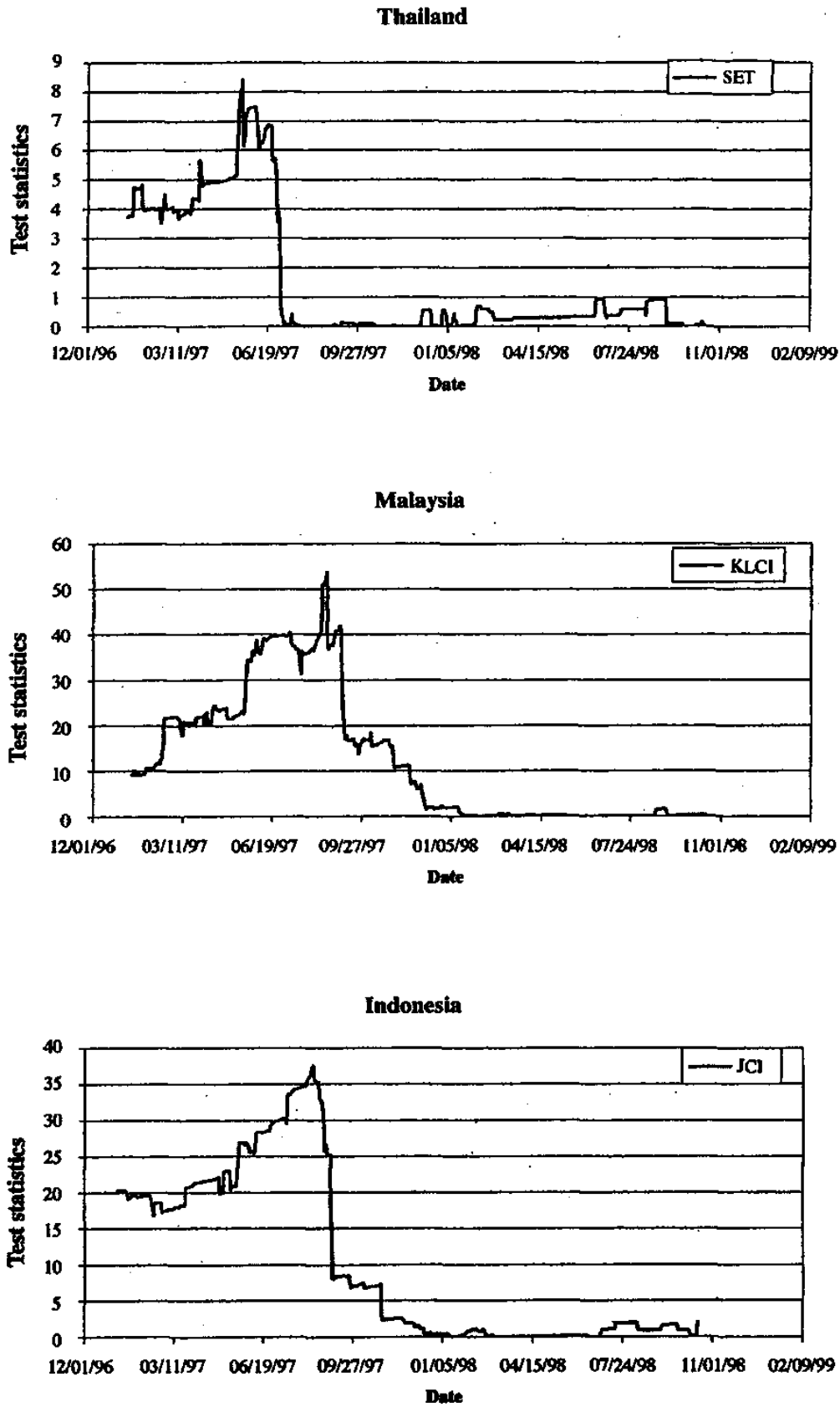


FIGURE 6
Rolling test (95% critical value = 2.25)

tests both in power and in estimation of the breakpoint. An alternative to the sequential test is therefore simply to treat the recursive and reverse recursive tests separately as one-sided tests.

When our methodology is applied to Asian stock market indexes the empirical evidence shows that both the recursive test and the rolling test reject the null of tail shape constancy for Malaysia and Indonesia. The breakpoints picked by both tests match well with suspected breakpoints arising from known changes in institutional arrangements. However, unlike the recursive test and due to the nature of its construction, the rolling test often does not produce a singleton supremum for the estimation of the break point. Therefore, separate one-sided tests of the recursive and reverse recursive tests seem preferable when the breakpoint date is of interest, as it is in this application.

APPENDIX PROOFS

Proof of Theorem 1. From Hall (1982, p. 41) and Hall-Welsh (1985, p. 338) we have for (i)

$$\alpha(\hat{\alpha}_t^{-1} - \alpha^{-1}) = m_t^{-1} \sum_{j=t-m_t+1}^t (Z_j' - 1) - dc^{-\rho} \rho(1 + \rho)^{-1} (m_t/w_t)^\rho + o_p(1) \quad (24)$$

where $\rho = \beta/\alpha$, Z_j' are independent exponential random variables with mean 1, and the $o_p(1)$ error holds uniformly in t . Set

$$m_t = [\kappa t] = [\kappa r T] = [lT]. \quad (25)$$

By Donsker's Theorem and with $t = [rT]$ we get

$$S_T = T^{-1/2} \sum_{j=t-m_t+1}^t (Z_j' - 1) \xrightarrow{d} W(l).$$

Let $\zeta = dc^{-\rho} \rho(1 + \rho)^{-1}$, so that we can write

$$\begin{aligned} \left(\frac{t}{T}\right)^{1/2} m_t^{1/2} \alpha(\hat{\alpha}_t^{-1} - \alpha^{-1}) &= \left(\frac{t}{m_t}\right)^{1/2} T^{-1/2} \sum_{j=t-m_t+1}^t (Z_j' - 1) - \zeta (m_t/w_t)^\rho m_t^{1/2} \left(\frac{t}{T}\right)^{1/2} \\ &= \left(\frac{t}{m_t}\right)^{1/2} S_T + o_p(1) \text{ (if } m_t < w_t^{\rho/(\rho+1/2)} = w_t^{2\beta/(2\beta+\alpha)}) \\ &\xrightarrow{d} \kappa^{-1/2} W(\kappa r) \\ &=_{d} W(r). \end{aligned} \quad (26)$$

Then, since $\alpha(\hat{\alpha}_t^{-1} - \alpha^{-1}) = -\hat{\alpha}_t^{-1}(\hat{\alpha}_t - \alpha)$, we have

$$\left(\frac{t}{T}\right)^{1/2} m_t^{1/2} \alpha^{-1}(\hat{\alpha}_t - \alpha) \xrightarrow{d} -W(r) =_{d} W(r)$$

for (A)'(i). For (B)'(i) if $m_t = \lambda w_t^{2\rho/(2\rho+1)}$ then

$$\zeta (m_t/w_t)^\rho m_t^{1/2} \left(\frac{t}{T}\right)^{1/2} = \zeta \lambda^{\rho+1/2} \left(\frac{t}{T}\right)^{1/2} \xrightarrow{p} \varphi r^{1/2},$$

as required.

For the rolling estimator, our subsample size is $w_t^* = [Tr] - [T(r - \gamma_0)]$. Then, instead of (25), set $m_{w_t^*} = [\kappa w_t^*] = [\kappa r T] - [\kappa s T]$. The results follow as before with

$$S_T^* = T^{-1/2} \sum_{j=t-m_{w_t^*}+1}^t (Z_j' - 1) \xrightarrow{d} W(\kappa r) - W(\kappa s) =_{d} \kappa^{1/2} W(r, \gamma_0).$$

For (B)'(ii) the noncentral term follows from $w_t^*/T \rightarrow r - s = \gamma_0$. The post-break sequential estimator follows by setting (25) as $m_{w_{2t}} = [\kappa w_{2t}] = [\kappa(1-r)T]$. ||

Proof of Theorem 2. The result for Q follows from

$$\begin{aligned} Y_T(t) &= \left(\frac{w_t m_t}{T}\right)^{1/2} \hat{\alpha}_t^{-1}(\hat{\alpha}_t - \alpha) - \left(\frac{w_t}{T}\right)^{1/2} \left(\frac{m_t}{m_T}\right)^{1/2} m_t^{1/2} \hat{\alpha}_T^{-1}(\hat{\alpha}_T - \alpha) \\ &\xrightarrow{d} W(r) - rW(1), \end{aligned}$$

since $(m_t/T)/(m_T/T) = ([\kappa r T]/T)/([\kappa T]/T) \rightarrow r$. For Q^* we have

$$V_T(t) = \left(\frac{w_t^* m_{w_t}}{T}\right)^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_t^* - \hat{\alpha}_T) - \left(\frac{w_t^*}{T}\right)^{1/2} \left(\frac{m_{w_t}}{m_T}\right)^{1/2} m_T^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_T - \alpha) \\ \xrightarrow{d} W(r, \gamma_0) - (r-s)W(1, 1),$$

since $(m_{w_t}/T)/(m_T/T) = ([\kappa r T] - [\kappa s T])/([\kappa T])/([\kappa T])/T \rightarrow r-s$. Similarly for Q^* , noting that $\hat{\alpha}_{2T}^{-1} \xrightarrow{d} \alpha$ under the null,

$$Z_T(t) = \left(\frac{tm_t}{T}\right)^{1/2} \hat{\alpha}_{2t}^{-1} (\hat{\alpha}_t - \alpha) - \left(\frac{t}{T}\right)^{1/2} \left(\frac{T}{w_{2t}}\right)^{1/2} \left(\frac{m_t}{m_{w_{2t}}}\right)^{1/2} \left[\left(\frac{w_{2t}}{T}\right)^{1/2} m_{w_{2t}}^{1/2} \hat{\alpha}_{2t}^{-1} (\hat{\alpha}_{2t} - \alpha)\right] \\ \xrightarrow{d} W(r) - \frac{r}{1-r} [W(1-r)],$$

and the result follows by the continuous mapping theorem. ||

Proof of Theorem 3. Let $\{X_i\}_{i=1}^{\tau-1}$ and $\{X_i\}_{i=\tau}^T$ have indices α_1 and α_2 , respectively, where $T_1 = \tau - 1$ and $T_2 = T - \tau + 1$. We work with the m_T largest order statistics of the full sample of which m_{T_1} and m_{T_2} are each taken from $Y_1 = [1, \tau)$ and $Y_2 = [\tau, T]$. We let $X_{(T-m_T+1)}^T \in Y_1$, say. Then, Hill's estimator is written as

$$\hat{\alpha}_T^{-1} = \frac{1}{m_T} \sum_{i=1}^{m_{T_1}} \log X_{(T_1-i+1)}^{T_1} - \log X_{(T_1-m_{T_1}+1)}^{T_1} + \frac{1}{m_T} \sum_{i=1}^{m_{T_2}} \log X_{(T_2-i+1)}^{T_2}.$$

Following Hall (1982, equation (3)) and Hall-Welsh (1985, equation (1.1)), we make the transformation $x \rightarrow x^{-1}$ so (1) becomes

$$F(x) = cx^\alpha(1 + dx^\beta + o(x^\beta)) \quad \text{as } x \downarrow 0 \tag{27}$$

and

$$\hat{\alpha}_T^{-1} = \left\{ \log X_{(m_{T_1}+1)}^{T_1} - \frac{1}{m_T} \sum_{i=1}^{m_{T_1}} \log X_{(i)}^{T_1} \right\} - \left\{ \frac{1}{m_T} \sum_{i=1}^{m_{T_2}} \log X_{(i)}^{T_2} \right\} = A - B. \tag{28}$$

The following lemma is used in the proof of Theorem 3. It generalizes (24), the representation of Hill's estimator in terms of standard exponential random variables, to the case of i.n.i.d. data. The result (24) is a special case of our lemma with $m_T = m_{T_1}$ and $m_{T_2} = 0$.

Lemma 10. For i.n.i.d. data such that $\{X_i\}_{i=1}^{\tau-1}$ and $\{X_i\}_{i=\tau}^T$ have indices α_1 and α_2 , respectively, we have

$$\hat{\alpha}_T^{-1} = \left(1 - \frac{m_{T_1}}{m_T}\right) \log c^{-1/\alpha_1} - \left(\frac{m_{T_2}}{m_T}\right) \log c^{-1/\alpha_2} \\ + \alpha_1^{-1} \frac{m_{T_1}}{m_T} m_{T_1}^{-1} \sum_{j=\tau_1-m_{T_1}+1}^{\tau_1} Z_j^{T_1} + \alpha_2^{-1} \frac{m_{T_2}}{m_T} m_{T_2}^{-1} \sum_{j=\tau_2-m_{T_2}+1}^{\tau_2} Z_j^{T_2} - \alpha_1^{-1} dc^{-\rho_1} \\ \times \left[1 - (1 + \rho_1)^{-1} \left(\frac{m_{T_1}}{m_T}\right) \right] \left(\frac{m_{T_1}}{T_1}\right)^{\rho_1} + \alpha_2^{-1} dc^{-\rho_2} (1 + \rho_2)^{-1} \left(\frac{m_{T_2}}{T_2}\right)^{\rho_2} \left(\frac{m_{T_2}}{m_T}\right) + o_p(1),$$

where $\rho_i = \beta_i/\alpha_i$ and $Z_j^{T_i}$ are independent exponential random variables with mean 1 constructed from the order statistics in period Y_i , $i = 1, 2$.

Proof of Lemma 10. We make use of the following representation for $X_{(i)}^T$ from Hall-Welsh (equations (3.9) and (3.10)):

$$\log X_{(i)}^T = \log c^{-1/\alpha_i} - \alpha_i^{-1} Y_{T(i)} - \alpha_i^{-1} dc^{-\rho_i} \exp(-\rho_i Y_{T(i)}) + o_p(1), \tag{29}$$

where $Y_{T(i)} = \sum_{j=1}^{T_i-i+1} Z_j^T/T_i - j + 1$. Working with (28) and applying (29) we have for A,

$$A = \left\{ \left(1 - \frac{m_{T_1}}{m_T}\right) \log c^{-1/\alpha_1} \right\} - \left\{ \alpha_1^{-1} (Y_{T_1(m_{T_1}+1)} - \frac{1}{m_T} \sum_{i=1}^{m_{T_1}} Y_{T_1(i)}) \right\} \\ - \left\{ \alpha_1^{-1} \left(dc^{-\rho_1} \exp(-\rho_1 Y_{T_1(m_{T_1}+1)}) - dc^{-\rho_1} \frac{1}{m_T} \sum_{i=1}^{m_{T_1}} \exp(-\rho_1 Y_{T_1(i)}) \right) \right\} \\ = A1 + A2 - A3. \tag{30}$$

For A2,

$$\begin{aligned}
 A2 &= \alpha_1^{-1} \frac{m_{T_1}}{m_T m_{T_1}} \frac{1}{\sum_{i=0}^{m_{T_1}-1} \sum_{j=1}^{T_1-i}} \frac{Z_j^{T_1}}{T_1-j+1} - \alpha_1^{-1} \sum_{j=1}^{T_1-m_{T_1}} \frac{Z_j^{T_1}}{T_1-j+1} \\
 &= \alpha_1^{-1} \left(\frac{m_{T_1}}{m_T} - 1 \right) \sum_{j=1}^{T_1-m_{T_1}} \frac{Z_j^{T_1}}{T_1-j+1} + \alpha_1^{-1} \frac{m_{T_1}}{m_T m_{T_1}} \sum_{j=T_1-m_{T_1}+1}^{T_1} Z_j^{T_1} \\
 &= o_p(m_{T_1}^{-1/2}) + \alpha_1^{-1} \frac{m_{T_1}}{m_T m_{T_1}} \sum_{j=T_1-m_{T_1}+1}^{T_1} Z_j^{T_1}, \tag{31}
 \end{aligned}$$

where the last line follows from Hall-Welsh equation (3.3). For A3,

$$\begin{aligned}
 A3 &= \alpha_1^{-1} dc^{-\rho_1} \left[\exp(-\rho_1 Y_{T_1(m_{T_1}, i)}) - \frac{m_{T_1}}{m_T m_{T_1}} \sum_{i=1}^{m_{T_1}} \exp\left(-\rho_1 \sum_{j=1}^{T_1-i+1} \frac{Z_j^{T_1}}{T_1-j+1}\right) \right] \\
 &= \alpha_1^{-1} dc^{-\rho_1} \left[\exp(-\rho_1 Y_{T_1(m_{T_1}, i)}) - \frac{m_{T_1}}{m_T m_{T_1}} \sum_{i=1}^{m_{T_1}} \right. \\
 &\quad \left. \times \exp(-\rho_1 Y_{T_1(m_{T_1}, i)}) \exp\left(-\rho_1 \sum_{j=T_1-m_{T_1}+1}^{T_1} \frac{Z_j^{T_1}}{T_1-j+1}\right) \right] \\
 &= \alpha_1^{-1} dc^{-\rho_1} \exp(-\rho_1 Y_{T_1(m_{T_1}, i)}) \left[1 - \frac{1}{m_{T_1}} \sum_{i=1}^{m_{T_1}} \exp\left(-\rho_1 \sum_{j=T_1-m_{T_1}+1}^{T_1} \frac{Z_j^{T_1}}{T_1-j+1}\right) \left(\frac{m_{T_1}}{m_T} \right) \right] \\
 &= \alpha_1^{-1} dc^{-\rho_1} \left(\frac{m_{T_1}}{T_1} \right)^{\rho_1} \left[1 - (1 + \rho_1)^{-1} \left(\frac{m_{T_1}}{m_T} \right) \right] + o_p(1), \tag{32}
 \end{aligned}$$

where the last line follows from Hall-Welsh Lemma 3.1. Thus we have

$$A = \left(1 - \frac{m_{T_1}}{m_T} \right) \log c^{-1/\alpha_1} + \alpha_1^{-1} \frac{1}{m_T} \sum_{j=T_1-m_{T_1}+1}^{T_1} Z_j^{T_1} - \alpha_1^{-1} dc^{-\rho_1} \left(\frac{m_{T_1}}{T_1} \right)^{\rho_1} \left[1 - (1 + \rho_1)^{-1} \left(\frac{m_{T_1}}{m_T} \right) \right] + o_p(1), \tag{33}$$

from (30), (31) and (32).

For B we have

$$\begin{aligned}
 B &= \frac{m_{T_2}}{m_T} \log c^{-1/\alpha_2} - \alpha_2^{-1} \frac{1}{m_T} \sum_{i=1}^{m_{T_2}} Y_{T_2(i)} - \alpha_2^{-1} dc^{-\rho_2} \frac{1}{m_T} \sum_{i=1}^{m_{T_2}} \exp(-\rho_2 Y_{T_2(i)}) + o_p(1) \\
 &= B1 - B2 - B3. \tag{34}
 \end{aligned}$$

For B2 we have

$$\begin{aligned}
 B2 &= \alpha_2^{-1} \frac{m_{T_2}}{m_T m_{T_2}} \frac{1}{\sum_{i=0}^{m_{T_2}-1} \sum_{j=1}^{T_2-i}} \frac{Z_j^{T_2}}{T_2-j+1} \\
 &= \alpha_2^{-1} \frac{m_{T_2}}{m_T} \left(\frac{1}{m_{T_2}} \sum_{j=T_2-m_{T_2}+1}^{T_2} Z_j^{T_2} + \sum_{j=1}^{T_2-m_{T_2}} \frac{Z_j^{T_2}}{T_2-j+1} \right) \\
 &= \alpha_2^{-1} \frac{m_{T_2}}{m_T} \left(\frac{1}{m_{T_2}} \sum_{j=T_2-m_{T_2}+1}^{T_2} Z_j^{T_2} + o_p(m_{T_2}^{-1/2}) \right)
 \end{aligned}$$

from Hall-Welsh equation (3.3) and for B3,

$$\begin{aligned}
 B3 &= \alpha_2^{-1} dc^{-\rho_2} \frac{m_{T_2}}{m_T m_{T_2}} \frac{1}{\sum_{i=1}^{m_{T_2}}} \exp\left(-\rho_2 \sum_{j=1}^{T_2-i+1} \frac{Z_j^{T_2}}{T_2-j+1}\right) \\
 &= \alpha_2^{-1} dc^{-\rho_2} \frac{m_{T_2}}{m_T m_{T_2}} \frac{1}{\sum_{i=0}^{m_{T_2}-1}} \exp(-\rho_2 Y_{T_2(m_{T_2}, i)}) \exp\left(-\rho_2 \sum_{j=T_2-m_{T_2}+1}^{T_2} \frac{Z_j^{T_2}}{T_2-j+1}\right) \\
 &= \alpha_2^{-1} dc^{-\rho_2} \frac{m_{T_2}}{m_T} \exp(-\rho_2 Y_{T_2(m_{T_2}, i)}) \left(\frac{1}{m_{T_2}} \sum_{i=0}^{m_{T_2}-1} \exp\left(-\rho_2 \sum_{j=T_2-m_{T_2}+1}^{T_2} \frac{Z_j^{T_2}}{T_2-j+1}\right) \right) \\
 &= \alpha_2^{-1} dc^{-\rho_2} \frac{m_{T_2}}{m_T} \left(\frac{m_{T_2}}{T_2} \right)^{\rho_2} (1 + \rho_2)^{-1} + o_p(1),
 \end{aligned}$$

from Hall–Welsh Lemma 3.1. Thus

$$B = \frac{m_{T_2}}{m_T} \log c^{-1/\alpha_2} - \alpha_2^{-1} \frac{1}{m_T} \sum_{j=T_2-m_{T_2}+1}^{T_2} Z_j^{T_2} - \alpha_2^{-1} d c^{-\rho_2} \frac{m_{T_2}}{m_T} \left(\frac{m_{T_2}}{T_2} \right)^{\rho_2} (1 + \rho_2)^{-1} + o_p(1), \quad (35)$$

and the lemma follows from (33) and (35). \square

The proof of Theorem 3 follows from Lemma 10. Let

$$\frac{m_{T_1}}{m_T} \rightarrow \pi_1 \in [0, 1] \quad \text{and} \quad \frac{m_{T_2}}{m_T} \rightarrow \pi_2 \in [0, 1].$$

A necessary condition for consistency of $\hat{\alpha}_t$ for α_t is that $m_{T_1} = o(T_1)$ and $m_{T_2} = o(T_2)$; a sufficient condition being

$$m_{T_1} = T^{\phi_1} \quad \text{and} \quad m_{T_2} = T^{\phi_2}, \quad (36)$$

with $\phi_i = 2\beta_i/(2\beta_i + \alpha_i) < 1$ for $i = 1, 2$. Then since

$$\frac{1}{m_{T_1}} \sum_{j=T_1-m_{T_1}+1}^{T_1} Z_j^{T_1} \xrightarrow{p} 1 \quad \text{and} \quad \frac{1}{m_{T_2}} \sum_{j=T_2-m_{T_2}+1}^{T_2} Z_j^{T_2} \xrightarrow{p} 1,$$

by the WLLN we have from Lemma 10,

$$\hat{\alpha}_T^{-1} \xrightarrow{p} \alpha_1^{-1} \pi_1 + \alpha_2^{-1} \pi_2 + (1 - \pi_1) \log c^{-1/\alpha_1} - \pi_2 \log c^{-1/\alpha_2} = \alpha^{-1}, \quad (37)$$

assuming (36) (or $m_{T_1} = o(T_1)$ and $m_{T_2} = o(T_2)$) holds.

We set $m_{T_i}/m_T = T^{\phi_i}/T^{\phi}$ in what follows with $\phi = (2\beta/2\beta + \alpha)$. We require $\alpha > 0$ so π_1 and/or π_2 have to be greater than zero. Suppose $\alpha_1 < \alpha_2$ then $T^{\phi_1} > T^{\phi_2}$ and

- (i) if $T^{\phi_1}/T^{\phi} \rightarrow \pi_1 > 0$ then $T^{\phi_2}/T^{\phi} \rightarrow \pi_2 = 0$ so $\pi_1 = 1$.
- (ii) if $T^{\phi_2}/T^{\phi} \rightarrow \pi_2 > 0$ then $T^{\phi_1}/T^{\phi} \rightarrow \infty$ which violates the requirement that $\pi_i \in [0, 1]$.

Thus, if $\alpha_1 < \alpha_2$ then $\hat{\alpha}_T^{-1} \xrightarrow{p} \alpha_1^{-1} = \alpha^{-1}$.

Suppose $\alpha_1 > \alpha_2$ then $T^{\phi_1} < T^{\phi_2}$ and

- (i) if $T^{\phi_1}/T^{\phi} \rightarrow \pi_1 > 0$ then $T^{\phi_2}/T^{\phi} \rightarrow \infty$ which violates the requirement that $\pi_2 \in [0, 1]$.
- (ii) if $T^{\phi_2}/T^{\phi} \rightarrow \pi_2 > 0$ then $T^{\phi_1}/T^{\phi} \rightarrow \pi_1 = 0$ so $\pi_2 = 1$.

Thus, if $\alpha_1 > \alpha_2$ then $\hat{\alpha}_T^{-1} \xrightarrow{p} \alpha_2^{-1} = \alpha^{-1}$. \square

Proof of Theorem 4. Write

$$\begin{aligned} m_{w_t}^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_t - \hat{\alpha}_T) &= m_{w_t}^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_t - \alpha_t) - \left(\frac{m_{w_t}}{m_T} \right)^{1/2} m_T^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_T - \alpha) + m_{w_t}^{1/2} \hat{\alpha}_T^{-1} (\alpha_t - \alpha) \\ &= O_p(1) + O_p(1) + I. \end{aligned} \quad (38)$$

For the recursive test we have $m_T = \lambda_T T^{2\beta/(2\beta + \alpha)}$ and $\alpha = \min(\alpha_1, \alpha_2)$. Suppose Case A1 then $\alpha = \alpha_2$ so for $t \in Y_2$ we have $I = o_p(1)$ and for $t \in Y_1$ we have $I = O_p(m_t^{1/2})$. For Case A2 we have $\alpha = \alpha_1$ and $\alpha_2 = \alpha_1$ so $I = o_p(1)$ for $t \in Y_1, Y_2$.

For the rolling test $w_t = w_t^*$ in (38). Then for Case A1 $I = O_p(m_{w_t^*}^{1/2})$ and for Case A2 we have $I = O_p(m_{w_t^*}^{1/2})$ for $t \in [r + w_t^*, T]$. The sequential test is the sum of the recursive and reverse recursive test so the convergence is reversed for the latter—i.e. the reverse recursive test is inconsistent for Case A1 but for Case A2 we have $m_t^{1/2} \hat{\alpha}_{2t}^{-1} (\alpha_t - \alpha_{2t}) = O_p(m_t^{1/2})$. \square

Proof of Corollary 5. The result follows from (B) and the continuous mapping theorem. For Q we have

$$\begin{aligned} Y_T(t) &= \left(\frac{w_t m_t}{T} \right)^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_t - \alpha) - \left(\frac{w_t}{T} \right)^{1/2} \left(\frac{m_t}{m_T} \right)^{1/2} m_T^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_T - \alpha) \\ &\xrightarrow{d} W(r) + \varphi r^{1/2} - r[W(1) + \varphi], \end{aligned}$$

and for Q^* ,

$$\begin{aligned} V_T(t) &= \left(\frac{w_t^* m_{w_t^*}}{T} \right)^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_t^* - \hat{\alpha}_T) - \left(\frac{w_t^*}{T} \right)^{1/2} \left(\frac{m_{w_t^*}}{m_T} \right)^{1/2} m_T^{1/2} \hat{\alpha}_T^{-1} (\hat{\alpha}_T - \alpha) \\ &\xrightarrow{d} W(r, \gamma_0) + \varphi \gamma_0^{1/2} - (r - s)[W(1, 1) + \varphi], \end{aligned}$$

and finally

$$Z_T(t) = \left(\frac{tm_t}{T}\right)^{1/2} \hat{\alpha}_{2T}^{-1}(\hat{\alpha}_t - \alpha) - \left(\frac{t}{T}\right)^{1/2} \left(\frac{T}{w_{2t}}\right)^{1/2} \left(\frac{m_t}{m_{w_{2t}}}\right)^{1/2} \\ \times \left[\left(\frac{w_{w_{2t}}}{T}\right)^{1/2} m_{w_{2t}}^{1/2} \hat{\alpha}_{2t}^{-1}(\hat{\alpha}_{2t} - \alpha) \right] \xrightarrow{d} W(r) + \varphi r^{1/2} - \frac{r}{1-r} [W(1-r) + \varphi(1-r)^{1/2}],$$

as required. Under (C), divergence comes from the second term in (26) since the first term is $O_p(1)$ and

$$\zeta\left(\frac{w_t}{T}\right)^{1/2} \left(\frac{m_w^{(2p+1)/2}}{w_t^p}\right) = \zeta\left(\frac{w_t}{T}\right)^{1/2} \left(\frac{m_{w_t}}{w_t^{2p/(2p+1)}}\right)^{(2p+1)/2} = \zeta\left(\frac{w_t}{T}\right)^{1/2} \left(\frac{m_{w_t}}{w_t^{2\beta/(2\beta+\alpha)}}\right)^{(2p+1)/2} \rightarrow \infty.$$

Proof of Theorem 6. The correction terms follow from the noncentral terms of Corollary 5 under (B). \parallel

Proof of Theorem 8. We write the process Y_t^T as Y_t in what follows to simplify notation. We use the following result from Resnick–Starica (1996, Lemma 4.1c): for a strictly stationary process $\{Y_t\}$ that satisfies (17) we have for $\varepsilon > 0$,

$$P(Y_1 > x, Y_t > y) \leq P(Y_0 > x)P(Y_0 > y - \varepsilon) + P(Y_0 > x, G_1^{-1} Y_0 > \varepsilon). \quad (39)$$

We also use the following fact extensively: for independent random variables η and Y_0 ,

$$P(\eta Y_0 > x) - E\eta^{\alpha/2} P(Y_0 > x) \quad (40)$$

from equation (2.10) of Mikosch–Starica (1999). Note that from regular variation of Y we have

$$P(Y_0 > x) - cx^{-\alpha/2} \quad \text{as } x \rightarrow \infty.$$

To show (a), for EC_T, C_T , we write

$$E\left(\log \frac{Y_1}{b}\right)_+ \left(\log \frac{Y_t}{b}\right)_+ \\ = \int_{x=1}^{\infty} \int_{y=1}^{\infty} P(Y_1 > bx, Y_t > by) \frac{dy}{y} \frac{dx}{x} \\ \leq \int_{x=1}^{\infty} \int_{y=1}^{\infty} P(Y_0 > bx) P(Y_0 > b(y - \varepsilon)) \frac{dy}{y} \frac{dx}{x} \\ + \int_{x=1}^{\infty} P(Y_0 > bx, G_1^{-1} Y_0 > b\varepsilon) \frac{dx}{x} \\ = \int_{x=1}^{\infty} \int_{y=1}^x P(Y_0 > bx) \frac{dy}{y} \frac{dx}{x} + \int_{y=1}^{\infty} \int_{x=1}^y P(Y_0 > b(y - \varepsilon)) \frac{dx}{x} \frac{dy}{y} \\ + \int_{x=1}^{\infty} P(Y_0 > bx, G_1^{-1} Y_0 > b\varepsilon) \frac{dx}{x} \\ = A1 + A2 + A3 \quad (41)$$

where the inequality follows from (39). For A1 we have

$$A1 = \int_{x=1}^{\infty} \int_{y=1}^x P(Y_0 > bx) \frac{dy}{y} \frac{dx}{x} \\ = \int_{x=1}^{\infty} \log x P(Y_0 > bx) \frac{dx}{x} \rightarrow \frac{m_T}{T} c \int_{x=1}^{\infty} \log x (x^{-\alpha/2}) \frac{dx}{x}, \quad (\text{as } x \rightarrow \infty) \\ = \frac{m_T}{T} c \int_{u=0}^{\infty} u e^{-u\alpha} du, \quad (u = \log x) \\ = \frac{m_T}{T} \frac{c}{\alpha^2}$$

and similarly for A2,

$$\begin{aligned} A2 &= \int_{y=1}^{\infty} \int_{x=1}^y P(Y_0 > b(y-\varepsilon)) \frac{dx}{x} \frac{dy}{y} \\ &= \int_{y=1}^{\infty} \log y P(Y_0 > b(y-\varepsilon)) \frac{dy}{y} \rightarrow \frac{m_T}{T} c \int_{y=1}^{\infty} \log y (y^{-\alpha}) \frac{dy}{y}, \quad (\text{as } y \rightarrow \infty, \varepsilon \rightarrow 0) \\ &= \frac{m_T}{T} \frac{c}{\alpha^2}. \end{aligned}$$

For A3 we use (40),

$$\begin{aligned} A3 &= \int_{x=1}^{\infty} P(Y_0 > bx, G_1^{\prime-1} Y_0 > b\varepsilon) \frac{dx}{x} \\ &= \int_{x=1}^{\infty} P((1 \wedge G_1^{\prime-1}) Y_0 > b(x \vee \varepsilon)) \frac{dx}{x} = E(G_1^{\prime-1})^\alpha \int_{x=1}^{\infty} P(Y_0 > b(x \vee \varepsilon)) \frac{dx}{x} \\ &\rightarrow \frac{m_T}{T} c E(G_1^{\prime-1})^\alpha \int_{x=1}^{\infty} x^{-\alpha} \frac{dx}{x} = \frac{m_T}{T} \frac{c}{\alpha} E(G_1^{\prime-1})^\alpha, \end{aligned}$$

and the result follows by substitution into (41). We assume $x \geq y$ for (b) and (c). For (b),

$$\begin{aligned} &E\left(\log \frac{Y_1}{b}\right)_+ I(\log Y_j > \log b + \delta/\sqrt{m_T}) \\ &= \int_{x=y}^{\infty} P(Y_1 > bx, Y_1 > be^{\delta/\sqrt{m_T}y}) \frac{dx}{x} \\ &\leq \int_{x=y}^{\infty} P(Y_0 > bx) P(Y_0 > be^{\delta/\sqrt{m_T}(y-\varepsilon)}) \frac{dx}{x} \\ &\quad + \int_{x=y}^{\infty} P(Y_0 > bx, G_1^{\prime-1} Y_0 > be^{\delta/\sqrt{m_T}\varepsilon}) \frac{dx}{x} \\ &\rightarrow \frac{m_T}{T} c \int_{x=y}^{\infty} x^{-\alpha} \frac{dx}{x} + \frac{m_T}{T} c E(G_1^{\prime-1})^\alpha \int_{x=y}^{\infty} x^{-\alpha} \frac{dx}{x} = \frac{m_T}{T} \frac{c}{\alpha} y^{-\alpha} (1 + E(G_1^{\prime-1})^\alpha) \end{aligned}$$

and

$$\begin{aligned} &EI(\log Y_1 > \log b + \delta/\sqrt{m_T}) \left(\log \frac{Y_j}{b}\right)_+ \\ &= \int_{y=0}^x P(Y_1 > be^{\delta/\sqrt{m_T}y}, Y_j > bx) \frac{dy}{y} \\ &\leq \int_{y=0}^x P(Y_0 > be^{\delta/\sqrt{m_T}y}) P(Y_0 > b(x-\varepsilon)) \frac{dy}{y} + \int_{y=0}^x P(Y_0 > be^{\delta/\sqrt{m_T}y}, G_1^{\prime-1} Y_0 > b\varepsilon) \frac{dy}{y} \\ &\rightarrow \frac{m_T}{T} c \int_{y=0}^x y^{-\alpha} \frac{dy}{y} + \frac{m_T}{T} c E(G_1^{\prime-1})^\alpha \int_{y=0}^x y^{-\alpha} \frac{dy}{y} = \frac{m_T}{T} \frac{c}{\alpha} x^{-\alpha} (1 + E(G_1^{\prime-1})^\alpha). \end{aligned}$$

Finally, we have

$$\begin{aligned} &E(I(\log Y_1 > \log b + \delta/\sqrt{m_T}), I(\log Y_j > \log b + \delta/\sqrt{m_T})) \\ &= P(Y_1 > be^{\delta/\sqrt{m_T}x}, Y_1 > be^{\delta/\sqrt{m_T}y}) \\ &\leq P(Y_0 > be^{\delta/\sqrt{m_T}x}) P(Y_0 > be^{\delta/\sqrt{m_T}(y-\varepsilon)}) \\ &\quad + P(Y_0 > be^{\delta/\sqrt{m_T}x}, G_1^{\prime-1} Y_0 > be^{\delta/\sqrt{m_T}\varepsilon}) \rightarrow \frac{m_T}{T} c x^{-\alpha} + \frac{m_T}{T} c E(G_1^{\prime-1})^\alpha x^{-\alpha}. \end{aligned}$$

for part (c).

For parts (d)–(f), note that the difference equation is written in terms of σ_t^2 . Thus regular variation of the tails of X_t^2 comes from regular variation of σ_t^2 . The proof of (d)–(f) follows (a)–(c) with the following modification. Note that C_T and D_T are defined in terms of X_t^2 so,

$$P(X_1^2 > x, X_T^2 > y) = E(Z_1^2)^{\alpha/2} P(\sigma_1^2 > x, \sigma_T^2 > y) = EZ_1^\alpha P(Y_1 > x, Y_T > y)$$

and (39) is applied to the last expression.

APPENDIX A

TABLE

Critical values for recursive, rolling and sequential tests

Tests	0.50	0.60	0.70	0.80	0.90	0.95	0.975	0.99
$Q(r)$	0.67	0.79	0.94	1.14	1.46	1.78	2.11	2.54
$Q^*(r)$								
$\gamma_0 = 0.15$	0.73	0.81	0.92	1.05	1.26	1.46	1.65	1.90
$\gamma_0 = 0.20$	0.82	0.93	1.06	1.23	1.49	1.75	2.01	2.30
$\gamma_0 = 0.25$	0.88	1.00	1.14	1.32	1.67	1.98	2.23	2.55
$\gamma_0 = 0.30$	0.90	1.04	1.21	1.43	1.78	2.12	2.45	2.86
$\gamma_0 = 0.35$	0.91	1.05	1.22	1.46	1.79	2.14	2.52	2.96
$\gamma_0 = 0.40$	0.89	1.04	1.20	1.44	1.84	2.23	2.57	3.09
$\gamma_0 = 0.45$	0.88	1.02	1.21	1.45	1.89	2.27	2.62	3.06
$\gamma_0 = 0.50$	0.89	1.04	1.23	1.47	1.87	2.25	2.64	3.11
$Q^*(r)$	4.62	5.73	7.20	9.48	13.98	18.31	22.94	28.82

Acknowledgements. Our thanks go to Rohit Deo at NYU for numerous discussions and valuable insights. We also thank Orazio Attanasio, the Editor, and three referees for extensive comments. Phillips thanks the NSF for research support under grant No. SBR 97-30295. This paper was circulated under the title "Measuring Change in Tail Behaviour with an Application to Emerging Financial Markets in Asia", NYU working paper series, SOR-99-2.

REFERENCES

- ANDREWS, D. W. K. (1993), "Tests for Parameter Stability and Structural Change with Unknown Change Point", *Econometrica*, **59**, 817–858.
- ANDREWS, D. W. K. and PLOBERGER, W. (1994), "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative", *Econometrica*, **62**, 1383–1414.
- BANERJEE, A., LUMSDAINE, L. and STOCK, J. (1992), "Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence", *Journal of Business and Economic Statistics*, **10**(3), 271–287.
- BOOTHE, P. and GLASSMAN, D. (1987), "The Statistical Distribution of Exchange Rates", *Journal of International Economics*, **22**, 297–319.
- DANIELSSON, J. and DE VRIES, C. G. (1997a), "Beyond the Sample: Extreme Quantile and Probability Estimation" (Manuscript, Tinbergen Institute, Rotterdam).
- DANIELSSON, J. and DE VRIES, C. G. (1997b), "Tail Index and Quantile Estimation with Very High Frequency Data", *Journal of Empirical Finance*, **4**, 241–257.
- DAVIS, R. D. and MIKOSCH, T. (1998), "The Sample ACF of Heavy-Tailed Stationary Processes with Applications to ARCH", *Annals of Statistics*, **26**(5), 2049–2080.
- DUMOUCHEL, W. H. (1983), "Estimating the Stable Index α in Order to Measure Tail Thickness: A Critique", *Annals of Statistics*, **11**, 1019–1031.
- EMBRECHTS, P., KLUPPELBERG, C. and MIKOSCH, T. (1997), *Modelling Extremal Events* (New York: Springer Verlag).
- HAEUSLER, E. and TEUGELS, J. L. (1985), "On Asymptotic Normality of Hill's Estimator for the Exponent of Regular Variation", *Annals of Statistics*, **13**(2), 743–756.
- HALL, P. (1982), "On Some Simple Estimates of an Exponent of Regular Variation", *Journal of the Royal Statistical Society B*, **44**(1), 37–42.
- HALL, P. and WELSH, A. H. (1985), "Adaptive Estimates of Parameters of Regular Variation", *The Annals of Statistics*, **13**(1), 331–341.
- HAWKINS, D. L. (1987), "A Test for a Change Point in a Parametric Model Based on a Maximal Wald-Type Statistic", *Sankhya*, **49**, 368–376.

- HILL, B. M. (1975), "A Simple General Approach to Inference about the Tail of a Distribution", *Annals of Statistics*, 3(5), 1163–1174.
- HSING, T. (1991), "On Tail Index Estimation Using Dependent Data", *Annals of Statistics*, 19(3), 1547–1569.
- HSING, T. (1993), "Extremal Index Estimation for a Weakly Dependent Stationary Sequence", *Annals of Statistics*, 21(4), 2043–2071.
- KEARNS, P. and PAGAN, A. (1997), "Estimating the Tail Density Tail Index for Financial Time Series", *Review of Economics and Statistics*, 79, 171–175.
- KOEDIJK, K. G. *et al.* (1990), "The Tail Index of Exchange Rate Returns", *Journal of International Economics*, 29, 93–108.
- LORETAN, M. and PHILLIPS, P. C. B. (1994), "Testing Covariance Stationarity of Heavy Tailed Time Series: An Overview of the Theory with Applications to Several Financial Datasets", *Journal of Empirical Finance*, 1, 211–248.
- MANDELBROT, B. B. (1963), "The Variation of Certain Speculative Prices", *Journal of Business*, 36, 394–419.
- MIKOSCH, T. and STARICA, C. (1998), "Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1, 1) Process" (working paper).
- PAGAN, A. R. and SCHWERT, G. W. (1990a), "Testing for Covariance Stationarity in Stock Market Data", *Economics Letters*, 33, 165–170.
- PAGAN, A. R. and SCHWERT, G. W. (1990b), "Alternative Models for Conditional Stock Volatility", *Journal of Econometrics*, 45, 267–290.
- PHILLIPS, P. C. B. and LORETAN, M. (1990), "Testing Covariance Stationarity under Moment Condition Failure with an Application to Common Stock Returns" (Cowles Foundation Discussion Paper No. 947).
- PHILLIPS, P. C. B. and LORETAN, M. (1994), "On the Theory of Testing Covariance Stationarity Under Moment Condition Failure", in G. S. Maddala, P. C. B. Phillips and T. N. Srinivasan (eds.), *Econometrics Festschrift for C. R. Rao* (Oxford: Basil Blackwell).
- PLOBERGER, W., KRAMER, W. and KONTRUS, K. (1989), "A New Test for Structural Stability in the Linear Regression Model", *Journal of Econometrics*, 40, 307–318.
- QUINTOS, C. E. (1999), "Tail Index Estimation and Value-at-Risk with Dependent Data" (New York University working paper).
- RESNICK, S. and STARICA, C. (1995), "Consistency of Hill's Estimator for Dependent Data", *J. Applied Probability*, 32, 129–167.
- RESNICK, S. and STARICA, C. (1996), "Tail Index Estimation for Dependent Data", (Cornell University, working paper).
- SAMORODNITSKY, G. and TAQQU, M. (1994), *Stable Non-Gaussian Random Processes* (New York: Chapman and Hall).