Statistica Sinica Preprint No: SS-2020-0265										
Title	Change-Point Tests for the Tail Parameter of Long									
	Memory Stochastic Volatility Time Series									
Manuscript ID	SS-2020-0265									
URL	http://www.stat.sinica.edu.tw/statistica/									
DOI	10.5705/ss.202020.0265									
Complete List of Authors	Annika Betken,									
	Davide Giraudo and									
	Rafal Kulik									
Corresponding Author	Davide Giraudo									
E-mail	davide.giraudo@rub.de									
Notice: Accepted version subject to English editing.										

Statistica Sinica: Newly accepted Paper (accepted author-version subject to English editing)

Statistica Sinica

Change-point tests for the tail parameter of Long Memory Stochastic Volatility time series

1

Annika Betken and Davide Giraudo and Rafał Kulik

University of Twente (The Netherlands)

Faculty of Mathematics, Ruhr-Universität Bochum (Germany) Department of Mathematics and Statistics, University of Ottawa (Canada)

Abstract:

We consider a change-point test based on the Hill estimator to test for structural changes in the tail index of Long Memory Stochastic Volatility time series. In order to determine the asymptotic distribution of the corresponding test statistic, we prove a uniform reduction principle for the tail empirical process in a twoparameter Skorohod space. It is shown that such a process displays a dichotomous behavior according to an interplay between the Hurst parameter, i.e., a parameter characterizing the dependence in the data, and the tail index. Our theoretical results are accompanied by simulation studies and the analysis of financial time series with regard to structural changes in the tail index.

Key words and phrases: stochastic volatility; long-range dependence; changepoint tests; tail empirical process; heavy tails; chaining

Introduction and motivation The tail behavior of the marginal 1. distribution of time series is of major relevance for statistics in applied sciences such as econometrics and hydrology, where heavy-tailed data occur frequently. More precisely, time series from finance such as the log returns of exchange rates and stock market indices display heavy tails; see Mandelbrot (1963). Furthermore, drastic events like the financial crisis in 2008 substantiate the importance of studying time series models that underlie financial data. Against this background, the identification of changes in the tail behavior of data-generating stochastic processes, that result in an increase or decrease in the probability of extreme events, is of utmost interest. In particular, the analysis of the tail behavior of financial data may pave the way for a corresponding adjustment of risk management for capital investments and, therefore, prevent huge capital losses. Indeed, there is empirical evidence that the tail behavior of financial time series may change over time: Quintos et al. (2001) identify changes in the tail of Asian stock market indices, Galbraith and Zernov (2004) find evidence for changes in the tail behavior of returns on U.S. equities, and Werner and Upper (2004) detect structural breaks in high-frequency data of Bund future returns.

2

1.1 Tail index estimation and change-point problem

Let X_j , $j \in \mathbb{N}$, be a stationary time series whose marginal tail distribution function \overline{F} is regularly varying with index $-\alpha$, $\alpha > 0$, i.e., $\mathbb{P}(X > x) = x^{-\alpha}L(x)$, where L is slowly varying at infinity. We recall that a measurable real-valued function is slowly varying at infinity if for all t > 0, $\lim_{x\to\infty} L(tx)/L(x) = 1$. Typical examples for L include constant functions or (iterated) logarithms. Since the tail behavior of X_j , $j \in \mathbb{N}$, is primarily determined by the value of the tail index α , identifying a change in the tail of data-generating processes corresponds to testing for a change-point in this parameter.

In particular, this means that, given a set of observations X_1, \ldots, X_n with $\mathbb{P}(X_j > x) = x^{-\alpha_j} L(x), j = 1, \ldots, n$, we aim at deciding on the testing problem (H, A) with

$$H: \alpha_1 = \dots = \alpha_n$$

and
$$A: \alpha_1 = \dots = \alpha_k \neq \alpha_{k+1} = \dots =$$

for some
$$k \in \{1, ..., n-1\}$$
.

 α_n

Test statistics that are designed for identifying structural changes in the tail index are naturally derived from an estimation of the tail index α . For

some general results on tail index estimation see Drees (1998a) and Drees (1998b). In this article, we focus on two estimators that are motivated by the fact that for a random variable X with tail index α

$$\lim_{u \to \infty} \mathbb{E}\left[\log\left(\frac{X}{u}\right) \mid X > u\right] = \lim_{u \to \infty} \frac{\mathbb{E}\left[\log\left(\frac{X}{u}\right) \mathbf{1}\left\{X > u\right\}\right]}{\mathbb{P}\left(X > u\right)} = \frac{1}{\alpha} =: \gamma$$

When we are given a set of observations X_1, \ldots, X_n , an approximation of the unknown distribution of X by its empirical analogue gives the following estimator for the tail index:

$$\widehat{\gamma} := \frac{1}{\sum_{j=1}^{n} \mathbf{1}\{X_j > u_n\}} \sum_{j=1}^{n} \log\left(\frac{X_j}{u_n}\right) \mathbf{1}\{X_j > u_n\} , \qquad (1.1)$$

where $u_n, n \in \mathbb{N}$, is a sequence with $u_n \to \infty$ and $n\bar{F}(u_n) \to \infty$. Replacing the deterministic levels u_n in the formula for $\hat{\gamma}$ by $X_{n:n-k_n}$ for some k_n , $1 \leq k_n \leq n-1$ such that $k_n \to \infty$, $k_n/n \to 0$, where $X_{n:n} \geq X_{n:n-1} \geq$ $\dots \geq X_{n:1}$ are the order statistics of the sample X_1, \dots, X_n , yields the Hill estimator

$$\widehat{\gamma}_{\text{Hill}} = \frac{1}{k_n} \sum_{i=1}^{k_n} \log\left(\frac{X_{n:n-i+1}}{X_{n:n-k_n}}\right).$$

As the most popular estimator for the tail index, established in Hill (1975), the Hill estimator has been widely studied in the literature. Its limiting distribution was obtained under various model assumptions, including linear processes (Resnick and Stărică (1997)), β -mixing processes (Drees (2000)), and Long Memory Stochastic Volatility models (Kulik and Soulier (2011)).

The first article that establishes a theory for change-point tests that are based on the Hill estimator seems to be Quintos et al. (2001). While Quintos et al. (2001) consider independent, identically distributed observations, ARCH- and GARCH-type processes, Kim and Lee (2011) and Kim and Lee (2012) extend their results to β -mixing processes and residual-based change-point tests for AR(p) processes with heavy-tailed innovations. In contrast, we study change-point tests for the tail index of Long Memory Stochastic Volatility time series based on the two estimators $\hat{\gamma}$ and $\hat{\gamma}_{\text{Hill}}$. In fact, our results are the first to consider the change-point problem for stochastic volatility models and time series with long-range dependence.

To motivate the design of test statistics for deciding on the change-point problem (H, A), we temporarily assume that the change-point location is known, i.e., for a given $k \in \{1, ..., n-1\}$ we consider the testing problem (H, A_k) with

$$A_k: \alpha_1 = \cdots = \alpha_k \neq \alpha_{k+1} = \cdots = \alpha_n.$$

For this testing problem, change-point tests have been considered in Phillips et al. (1990) and Koedijk et al. (1990). In order to decide on (H, A_k) , we compare an estimator $\hat{\gamma}_k$ of the tail index based on the observations X_1, \ldots, X_k to an estimator $\hat{\gamma}_n$ of the tail index based on the whole sample

 X_1, \ldots, X_n . This idea leads to studying the following test statistic

$$\Gamma_{k,n} = \frac{k}{n} \left| \frac{\widehat{\gamma}_k}{\widehat{\gamma}_n} - 1 \right|.$$

Under the assumption that the change-point location is unknown under the alternative, it seems natural to consider the statistic $\Gamma_{k,n}$ for every potential change-point location k and to decide in favor of the alternative hypothesis A if the maximum of its values exceeds a predefined threshold. As a result, a change-point test for the testing problem (H, A) that rests upon the estimator $\hat{\gamma}$ defined by (1.1) bases test decisions on the values of the statistic

$$\Gamma_n := \sup_{t \in [t_0, 1]} t \left| \frac{\widehat{\gamma}_{\lfloor nt \rfloor}}{\widehat{\gamma}_n} - 1 \right|$$
(1.2)

with $t_0 \in (0, 1)$ and with the sequential version of $\widehat{\gamma}$ defined by

$$\widehat{\gamma}_{\lfloor nt \rfloor} := \frac{1}{\sum_{j=1}^{\lfloor nt \rfloor} \mathbf{1}\{X_j > u_n\}} \sum_{j=1}^{\lfloor nt \rfloor} \log\left(\frac{X_j}{u_n}\right) \mathbf{1}\{X_j > u_n\}.$$
(1.3)

Likewise, a test statistic based on the Hill estimator is given by

$$\widetilde{\Gamma}_n := \sup_{t \in [t_0, 1]} t \left| \frac{\widehat{\gamma}_{\text{Hill}}(t)}{\widehat{\gamma}_{\text{Hill}}(1)} - 1 \right|$$

with the sequential version of $\widehat{\gamma}_{\text{Hill}}$ defined by

$$\widehat{\gamma}_{\mathrm{Hill}}(t) := \frac{1}{\lfloor k_n t \rfloor} \sum_{i=1}^{\lfloor k_n t \rfloor} \log \left(\frac{X_{\lfloor nt \rfloor : \lfloor nt \rfloor - i + 1}}{X_{\lfloor nt \rfloor : \lfloor nt \rfloor - k_{\lfloor nt \rfloor}}} \right).$$

In this context, the most comprehensive theory for change-point tests is presented in Hoga (2017). The author considers a number of test statistics based on different tail index estimators and derives their asymptotic distributions under the assumption of β -mixing data generating processes.

In the following, we derive the asymptotic distribution of both estimators, i.e., $\hat{\gamma}_{\lfloor nt \rfloor}$ and $\hat{\gamma}_{\text{Hill}}(t)$, and the corresponding tests statistics, i.e., Γ_n and $\tilde{\Gamma}_n$, under the hypothesis of stationary time series data. For this purpose, we first prove a limit theorem for the tail empirical process of Long Memory Stochastic Volatility time series in two parameters. This limit theorem does not necessarily relate to the change-point context. It can therefore be considered of independent interest and, thus, as the main theoretical result of our work. Our theoretical results are accompanied by simulation studies. As an empirical application of our tests, we consider Standard & Poor's 500 daily closing index covering the period from January 2008 to December 2008, the year of the financial crisis. We identify a change in the data at exactly one day after Lehman Brothers filed for bankruptcy protection, an event which is thought to have played a major role in the unfolding of the crisis in 2007 — 2008.

1.2 Tail empirical process

In order to derive the limit distribution of the tail estimators $\widehat{\gamma}_{\lfloor nt \rfloor}$ and $\widehat{\gamma}_{\rm Hill}(t)$, parametrized in t, and the corresponding test statistics Γ_n and $\widetilde{\Gamma}_n$, it is crucial to note that

$$\widehat{\gamma}_{\lfloor nt \rfloor} = \frac{1}{\sum_{j=1}^{\lfloor nt \rfloor} \mathbf{1}\{X_j > u_n\}} \sum_{j=1}^{\lfloor nt \rfloor} \log\left(\frac{X_j}{u_n}\right) \mathbf{1}\{X_j > u_n\}$$
$$= \frac{1}{\widetilde{T}_n(1,t)} \int_1^\infty s^{-1} \widetilde{T}_n(s,t) ds , \qquad (1.4)$$

where

$$\widetilde{T}_n(s,t) = \frac{1}{n\overline{F}(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \mathbf{1} \left\{ X_j > u_n s \right\}.$$

As a result, asymptotics of the considered statistics can be derived from a limit theorem for the two-parameter tail empirical process

$$e_n(s,t) := \left\{ \widetilde{T}_n(s,t) - T(s,t) \right\}, \ s \in [1,\infty], \ t \in [0,1],$$
(1.5)

where T(s,t) does not correspond to the mean of $\widetilde{T}_n(s,t)$, but rather to the limit of that mean, i.e., to

$$T(s,t) := ts^{-\alpha}.$$
(1.6)

Among others, the tail empirical process in one parameter, i.e., $e_n(s, 1)$, $s \in [1, \infty]$, has previously been studied in Mason (1988), Einmahl (1990), and Einmahl (1992) for independent, identically distributed observations,

in Rootzén (2009) for absolutely regular processes, and in Kulik and Soulier (2011) for Long Memory Stochastic Volatility time series. For the latter, the convergence of the two-parameter tail empirical process will be discussed in Section 2.2.

1.3 Long Memory Stochastic Volatility model

A phenomenon that is often encountered in the context of financial time series corresponds to the fact that the observations seem to be uncorrelated, whereas their absolute values or higher moments tend to be highly correlated. Another characteristic of financial time series is the occurrence of heavy tails. In particular, the distribution of the considered data often exhibits tails that are heavier than those of a normal distribution. The previously described features of financial data can be covered by stochastic volatility models.

Stochastic volatility model

The Long Memory Stochastic Volatility model that is taken as a basis of the theoretical results established in this article can be considered as a generalization of stochastic volatility models considered, for example, in Taylor (1986). Initially, this model had been introduced by Breidt et al.

(1998) and, independently, by Harvey (2002). An overview of stochastic volatility models with long-range dependence and their basic properties is given in Deo et al. (2006) and in Hurvich and Soulier (2009).

Stochastic volatility time series $X_j, j \in \mathbb{N}$, are typically defined via

$$X_j = Z_j \varepsilon_j$$
 with $Z_j = \exp\left(\frac{1}{2}Y_j\right)$, (1.7)

where $\varepsilon_j, j \in \mathbb{N}$, is a sequence of independent, identically distributed random variables with mean 0, and $Y_j, j \in \mathbb{N}$, is a Gaussian process, independent of $\varepsilon_j, j \in \mathbb{N}$.

While these models are often restricted to modeling a relatively fast decay of dependence in Y_j , $j \in \mathbb{N}$, the so-called Long Memory Stochastic Volatility model allows for long-range dependence. In what follows, we will specify a corresponding dependence structure for Y_j , $j \in \mathbb{N}$.

Transformed Gaussian processes

The rate of decay of the autocovariance function is crucial to the definition of long-range dependence in time series.

Definition 1. A (second-order) stationary, real-valued time series Y_j , $j \in \mathbb{Z}$, is called long-range dependent if its autocovariance function γ satisfies

$$\gamma_Y(k) := \operatorname{Cov}(Y_1, Y_{k+1}) \sim k^{-D} L_\gamma(k), \text{ as } k \to \infty,$$

with $D \in (0, 1)$ for a slowly varying function L_{γ} . We refer to D as longrange dependence (LRD) parameter; see Pipiras and Taqqu (2017), p. 17.

The transformed random variables $Z_j = G(Y_j), j \in \mathbb{N}$, can be considered as elements of the Hilbert space $L^2 := L^2(\mathbb{R}, \varphi(x)dx)$, i.e., the space of all measurable, real-valued functions which are square-integrable with respect to the measure $\varphi(x)dx$ associated with the standard normal density function φ , equipped with the inner product

$$\langle G_1, G_2 \rangle_{L^2} := \int_{-\infty}^{\infty} G_1(x) G_2(x) \varphi(x) dx = \mathbb{E} \left[G_1(Y) G_2(Y) \right]$$

where $G_1, G_2 \in L^2(\mathbb{R}, \varphi(x)dx)$ and Y denotes a standard normally distributed random variable. In order to characterize the dependence structure of transformed Gaussian processes, we consider their expansion in Hermite polynomials.

Definition 2. For $n \ge 0$, the Hermite polynomial of order n is defined by

$$H_n(x) = (-1)^n \mathrm{e}^{\frac{1}{2}x^2} \frac{d^n}{dx^n} \mathrm{e}^{-\frac{1}{2}x^2}, \ x \in \mathbb{R}.$$

The sequence of Hermite polynomials constitutes an orthogonal basis of L^2 . As a result, every $G \in L^2(\mathbb{R}, \varphi(x)dx)$ has an expansion in Hermite polynomials, i.e., for $G \in L^2(\mathbb{R}, \varphi(x)dx)$ and Y standard normally distributed,

we have

$$G(Y) \stackrel{L^2}{=} \sum_{r=0}^{\infty} \frac{J_r(G)}{r!} H_r(Y),$$
(1.8)

where $\|\cdot\|_{L^2}$ denotes the norm induced by the inner product $\langle\cdot,\cdot\rangle_{L^2}$.

Under the assumption that, as k tends to ∞ , $\gamma_Y(k)$ converges to 0 with a certain rate, the asymptotically dominating term in the series (1.8) is the summand corresponding to the smallest integer r for which the Hermite coefficient $J_r(G)$ is non-zero. This index, which decisively depends on G, is called Hermite rank.

Definition 3. Let $G \in L^2(\mathbb{R}, \varphi(x)dx)$, $\mathbb{E}[G(Y)] = 0$ for standard normally distributed Y and let $J_r(G)$, $r \ge 1$, be the Hermite coefficients in the Hermite expansion of G. The smallest index $k \ge 1$ for which $J_k(G) \ne 0$ is called the Hermite rank of G, i.e.,

$$r := \min \left\{ k \ge 1 : J_k(G) \neq 0 \right\}.$$

Given the previous definitions, we specify model assumptions that are taken as a basis for the results in the following sections.

Definition 4. Let the data generating process X_j , $j \in \mathbb{N}$, satisfy

$$X_j = Z_j \varepsilon_j, \quad j \in \mathbb{N},$$

where ε_j , $j \in \mathbb{N}$, is a sequence of independent, identically distributed random variables with mean 0, and Z_j , $j \in \mathbb{N}$, is a long-range dependent transformed Gaussian process with $Z_j = \sigma(Y_j)$, $j \in \mathbb{N}$, for some stationary, long-range dependent Gaussian process Y_j , $j \in \mathbb{N}$, with LRD parameter Dand a positive function σ . More precisely, assume that Y_j , $j \in \mathbb{N}$, admits a linear representation with respect to an independent, standard normally distributed sequence η_k , $k \in \mathbb{Z}$, i.e.,

$$Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}, \quad j \in \mathbb{N},$$

with $\sum_{k=1}^{\infty} c_k^2 = 1$. Furthermore, suppose that $(\varepsilon_j, \eta_j), j \in \mathbb{Z}$, is a sequence of independent, identically distributed random vectors. A sequence of random variables $X_j, j \in \mathbb{N}$, which satisfies the previous assumption is called a Long Memory Stochastic Volatility (LMSV) time series.

Remark 1. The model assumptions generalize the preceding concepts of stochastic volatility models with long-range dependence by allowing for general transformed Gaussian sequences Z_j , $j \in \mathbb{N}$, and dependence between the sequences $\{Y_j, j \in \mathbb{N}\}$, and $\{\varepsilon_j, j \in \mathbb{N}\}$. Instead of claiming mutual independence of Y_j , $j \in \mathbb{N}$, and ε_j , $j \in \mathbb{N}$, the sequence of random vectors (η_j, ε_j) is assumed to be independent. In particular, this implies that for a fixed index j, the random variables Y_j and ε_j are independent, while Y_j

may depend on ε_i , i < j. In many cases, an LMSV model incorporating this dependence structure is referred to as <u>LMSV with leverage</u>, as it allows for so-called <u>leverage effects</u> in financial time series. Not taking account of leverage, Definition 4 corresponds to the LMSV model considered in Kulik and Soulier (2011), while a similar model with leverage is considered in Bilayi-Biakana et al. (2019).

It can be shown that random variables X_j , $j \in \mathbb{N}$, satisfying Definition 4 are uncorrelated, while their squares inherit the dependence structure from the transformed Gaussian sequence Z_j^2 , $j \in \mathbb{N}$. Moreover, X_j , $j \in \mathbb{N}$, inherits the tail behavior from the sequence ε_j , $j \in \mathbb{N}$, if the marginal distribution of the random variables ε_j , $j \in \mathbb{N}$, has a regularly varying right tail, i.e., $\overline{F}_{\varepsilon}(x) := \mathbb{P}(\varepsilon_1 > x) = x^{-\alpha}L(x)$ for some $\alpha > 0$ and a slowly varying function L, and if $\mathbb{E}[\sigma^{\alpha+\delta}(Y_1)] < \infty$ for some $\delta > 0$. More precisely, under these assumptions the following asymptotic equivalence holds:

$$\mathbb{P}(X_1 > x) \sim \mathbb{E}[\sigma^{\alpha}(Y_1)] \mathbb{P}(\varepsilon_1 > x), \text{ as } x \to \infty.$$

This result is known as Breiman's Lemma; see Breiman (1965). On this account, it follows that Definition 4 is suited for modeling the previously described characteristic features of financial time series. In all following sections, we will therefore assume that the data-generating process X_j , $j \in \mathbb{N}$, corresponds to a LMSV time series specified by Definition 4.

1.4 Organisation of the paper

Equipped with the introductory remarks and definitions, we are in a position to discuss the structure of the paper. In Section 2 we state the technical assumptions that are needed for our theoretical results. These are followed by the main theorem on convergence of the two-parameter tail empirical process (Theorem 3). Convergence of estimators of the tail index (Corollary 1) and the test statistics (Corollary 2) are immediate consequences. Simulation studies are presented in Section 3, while real-data analysis can be found in Section 4. All the proofs are included in the supplementary document. In order to establish convergence of the two-parameter tail empirical process, we decompose it into a martingale and a long-range dependent part. The latter is dealt with in the supplementary document. For the former, we establish finite dimensional convergence using classical tools from martingale theory, while tightness of the two-parameter martingale part is handled by chaining. This is a theoretical novelty in the present context since the methods used in related papers are not suitable (the method used in Kulik and Soulier (2011) cannot be applied to models with leverage, while the approach in Bilayi-Biakana et al. (2019) is not well-suited for two-parameter processes).

2. Main results

2.1 Assumptions

In this section, we establish the assumptions guaranteeing convergence of the two-parameter tail empirical process for LMSV time series. Initially, we specify the LMSV model yielding the main assumptions for the theory:

Assumption 1 (Main Assumptions). Let $X_j = Z_j \varepsilon_j$, $j \in \mathbb{N}$, satisfy Definition 4 with $Z_j = \sigma(Y_j)$, $j \in \mathbb{N}$, for some stationary, long-range dependent Gaussian process Y_j , $j \in \mathbb{N}$, with autocovariance function $\gamma_Y(k) :=$ $\operatorname{Cov}(Y_1, Y_{k+1}) \sim k^{-D} L_{\gamma}(k)$, as $k \to \infty$, $D \in (0, 1)$, and some independent, identically distributed sequence ε_j , $j \in \mathbb{N}$, with regularly varying right tail, i.e., $\overline{F}_{\varepsilon}(x) := \mathbb{P}(\varepsilon_1 > x) = x^{-\alpha} L(x)$ for some $\alpha > 0$ and a slowly varying function L. Moreover, let r denote the Hermite rank of $\Psi(y) := \sigma^{\alpha}(y)$ and assume that r < 1/D.

We note that for a very strong dependence (D close to 0) a large range of the Hermite ranks is allowed. For D close to 1 only rank 1 is allowed.

In the following, we list some technical conditions that characterize the behavior of the slowly varying function L and the moments of $\sigma(Y_1)$. For this, we introduce another condition on the distribution function F_{ε} . This definition stems from Drees (1998c).

Definition 5 (Second order regular variation). Let $\overline{F}_{\varepsilon}(x) = x^{-\alpha}L(x)$ for some $\alpha > 0$ and some slowly varying function L that is represented by

$$L(x) = c \exp\left(\int_{1}^{x} \frac{\eta(u)}{u} du\right)$$

for some constant c and a measurable function η . Furthermore, we assume that there exists a bounded, decreasing function η^* on $[0, \infty)$, regularly varying at infinity with parameter $\rho \ge 0$, i.e., $\eta^*(x) = x^{-\rho}L_{\eta^*}(x)$, such that

$$|\eta(s)| \leqslant C\eta^*(s),$$

for some constant C and for all $s \ge 0$. We say that $\overline{F}_{\varepsilon}$ is second order regularly varying with tail index α and rate function η^* and we write $\overline{F}_{\varepsilon} \in 2\text{RV}(\alpha, \eta^*)$.

Second-order regular variation allows to control the difference between $\overline{F}_{\varepsilon}$ and the function $u \mapsto u^{-\alpha}$; see Lemma 1 and 2 in the supplementary file. Moreover, the specific form of L guarantees continuity of $\overline{F}_{\varepsilon}$.

Assumption 2 (Technical Assumptions). Suppose the main assumptions hold. Additionally, we assume that

(TA.1) $\overline{F}_{\varepsilon} \in 2\text{RV}(\alpha, \eta^*)$ and η is regularly varying with index ρ ;

(TA.2)
$$u_n \to \infty, \ n\overline{F}(u_n) \to \infty, \ \eta^*(u_n) = o\left(\frac{d_{n,r}}{n} + \frac{1}{\sqrt{n\overline{F}(u_n)}}\right)$$
, where $d_{n,r}$ is

2.1 Assumptions18

defined by

$$d_{n,r}^{2} = \operatorname{Var}\left(\sum_{j=1}^{n} H_{r}(Y_{j})\right) \sim c_{r} n^{2-rD} L_{\gamma}^{r}(n), \ c_{r} = \frac{2r!}{(1-Dr)(2-Dr)};$$
(2.9)

(TA.3) $\mathbb{E}\left[\sigma^{\alpha+\max\{\rho,\alpha\}+\vartheta}\left(Y_{1}\right)\right]<\infty$ for some $\vartheta>0;$

(TA.4) $\mathbb{E}\left[\left(\sigma\left(Y_{1}\right)\right)^{-1}\right] < \infty.$

Remark 2. Assumption (TA.2) handles the bias which is created by centering the tail empirical process not by its mean, but rather by the limit of that mean.

Example 1. The most commonly used second order assumption is that

$$L(x) = c \exp\left(\int_{1}^{x} \frac{\eta(u)}{u} du\right)$$

with $\eta(s) = s^{-\alpha\beta}$ for some $\beta > 0$. It then holds that $\overline{F}_{\varepsilon}(s) = C\left(s^{-\alpha} + \mathcal{O}(s^{-(\alpha(\beta+1))})\right)$, for $s \to \infty$, for some constant c > 0. Furthermore, we have

$$\sup_{s \ge 1} \left| \frac{\overline{F}_{\varepsilon}(u_n s)}{\overline{F}_{\varepsilon}(u_n)} - s^{-\alpha} \right| = \mathcal{O}(u_n^{-\alpha\beta}).$$

In this case, (TA.2) can be replaced by the assumption $u_n^{-\alpha\beta} = o\left(\frac{d_{n,r}}{n} + \frac{1}{\sqrt{nF(u_n)}}\right).$

2.2 Convergence of the tail empirical process19

2.2 Convergence of the tail empirical process

Recall that the tail empirical process in two parameters is defined by

$$e_n(s,t) := \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{\lfloor nt \rfloor} \mathbf{1} \{ X_j > u_n s \} - t s^{-\alpha}, \ s \in [1,\infty], \ t \in [0,1].$$

The following theorem establishes a characterization of its limit. In order to state this, we recall that a Hermite-Rosenblatt process of order rwith a self-similarity parameter H is a stochastic process $Z_{r,H}(t)$ defined for all $t \ge 0$ by a multiple Wiener-Ito integral with respect to a standard Brownian motion:

$$Z_{r,H}(t) = \omega(r,H) \int_{-\infty}^{\infty} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_{r-1}} \left(\int_0^t \prod_{j=1}^r (s-x_i)_+^{H-\frac{3}{2}} ds \right) dB(x_m) \cdots dB(x_1),$$

where $x_+ := \max(0, x)$ and $\omega(r, H) > 0$ satisfies

$$\omega^{2}(r,H) = \frac{r!(2r(H-1)+1)(r(H-1)+1)}{\left(\int_{0}^{\infty} [x(x+1)]^{H-\frac{3}{2}} dx\right)^{r}};$$

see Beran (2013), Section 3.7. In our case, $H = 1 - \frac{rD}{2}$ and hence the restriction r < 1/D gives H > 1/2. We recall that the standard Brownian motion has self-similarity parameter H = 1/2. Thus, H > 1/2 indeed indicates a presence of long memory.

Theorem 3. Let X_j , $j \in \mathbb{N}$, be a stationary time series with marginal tail distribution function \overline{F} . Moreover, assume that Assumptions 1 and 2 hold.

2.2 Convergence of the tail empirical process20

1. If
$$\frac{n}{d_{n,r}} = o\left(\sqrt{n\overline{F}(u_n)}\right)$$
, then as $n \to \infty$,
$$\frac{n}{d_{n,r}}e_n(s,t) \Rightarrow \frac{s^{-\alpha}}{\mathbb{E}\left[\sigma^{\alpha}(Y_1)\right]} \frac{J_r(\Psi)}{r!} Z_{r,H}(t), \qquad (2.10)$$

where $\Psi(y) = \sigma^{\alpha}(y)$, r is the Hermite rank of Ψ , $Z_{r,H}$ is an r-th order Hermite process, $H = 1 - \frac{rD}{2}$, and $d_{n,r}^2$ is defined in (2.9).

2. If
$$\sqrt{n\overline{F}(u_n)} = o\left(\frac{n}{d_{n,r}}\right)$$
, then as $n \to \infty$,
 $\sqrt{n\overline{F}(u_n)}e_n(s,t) \Rightarrow W(s^{-\alpha},t),$ (2.11)

where W denotes a standard Brownian sheet.

The convergence holds in a two-parameter Skorohod space, i.e., \Rightarrow denotes weak convergence in $D([1,\infty] \times [0,1])$.

The dichotomy of the limiting process is explained by the decomposition of the tail empirical process into the sum of a martingale and a partial sum of long-range dependent random variables, which can be viewed as a special case of Doob's decomposition; see the supplementary document. If $\frac{n}{d_{n,r}} = o\left(\sqrt{nF(u_n)}\right)$, the martingale part in the decomposition becomes negligible, such that the limiting process arises from the convergence of the long-range dependent part. If $\sqrt{nF(u_n)} = o\left(\frac{n}{d_{n,r}}\right)$, the long-range dependent part in the decomposition becomes negligible, such that the limiting

2.2 Convergence of the tail empirical process21

process arises from the convergence of the martingale part. The same decomposition has already been employed in Kulik and Soulier (2011), Betken and Kulik (2019), and Bilayi-Biakana et al. (2019).

The assumption $\sqrt{nF(u_n)} = o\left(\frac{n}{d_{n,r}}\right)$ yields the "standard" convergence (2.11), which in turn will imply the "standard" convergence for the change-point statistics studied below. This is important from an application point of view. Indeed, the quantiles of (functionals of) the limiting process in (2.11) are rather easy to simulate under the null hypothesis of no change. Furthermore, under this limiting regime, a validity of the i.i.d. bootstrap can be conjectured. On the other hand, the limiting process in (2.10) is much harder to simulate. Indeed, first, one has to know the Hermite rank and the α -th moment of the <u>unobservable</u> process Y_j . Next, even if we know this, simulation of the Hermite-Rosenblatt process is not an easy task. As such, it is important from a statistical point of view to be able, if possible, to work under the regime that guarantees the validity of (2.11). This is in fact the way we approach simulation studies.

Ignoring the slowly varying components, $\sqrt{nF(u_n)} = o\left(\frac{n}{d_{n,r}}\right)$ means that $n^{3/2-rD} = o(u_n^{\alpha/2})$. For a given long memory parameter D and the Hermite rank r, this induces restrictions on the thresholds u_n for the "standard" convergence (2.11) to hold. The stronger the memory is (that is, the smaller D is), the larger the threshold needed is. Intuitively, under strong dependence, we can use only very extreme observations in order to remove the effect of long memory.

Furthermore, if the rates $\frac{n}{d_{n,r}}$ and $\sqrt{n\overline{F}(u_n)}$ are asymptotically equivalent (up to a constant), it can be conjectured that a limiting process is a linear combination of (uncorrelated, but not independent) limiting processes that appear on the right-hand side of both (2.10)-(2.11).

2.3 Convergence of the tail estimators

Recall that the considered tail index estimators of $\gamma = 1/\alpha$ are defined by

$$\widehat{\gamma}_{\lfloor nt \rfloor} := \frac{1}{\sum_{j=1}^{\lfloor nt \rfloor} \mathbf{1}\{X_j > u_n\}} \sum_{j=1}^{\lfloor nt \rfloor} \log\left(\frac{X_j}{u_n}\right) \mathbf{1}\{X_j > u_n\}$$

and

$$\widehat{\gamma}_{\text{Hill}}(t) := \frac{1}{\lfloor k_n t \rfloor} \sum_{i=1}^{\lfloor k_n t \rfloor} \log\left(\frac{X_{\lfloor nt \rfloor : \lfloor nt \rfloor - i + 1}}{X_{\lfloor nt \rfloor : \lfloor nt \rfloor - k_{\lfloor nt \rfloor}}}\right)$$

where k_n and u_n are related by $k_n = \lfloor n\overline{F}(u_n) \rfloor$, so that $k_n \to \infty$ and $k_n/n \to 0$.

Based on Theorem 3 the limiting distributions of $\widehat{\gamma}_{\lfloor nt \rfloor}$ and $\widehat{\gamma}_{\mathrm{Hill}}(t)$ can be established in $D[t_0, 1]$ for any $t_0 \in (0, 1)$.

Corollary 1. Let X_j , $j \in \mathbb{N}$, be a stationary time series with marginal tail distribution function \overline{F} . Moreover, assume that Assumptions 1 and 2 hold.

2.3 Convergence of the tail estimators23

1. If
$$\frac{n}{d_{n,r}} = o\left(\sqrt{n\overline{F}(u_n)}\right)$$
, then as $n \to \infty$,
$$\frac{n}{d_{n,r}}t\left(\widehat{\gamma}_{\lfloor nt \rfloor} - \gamma\right) \Rightarrow 0 , \quad \frac{n}{d_{n,r}}t\left(\widehat{\gamma}_{\mathrm{Hill}}(t) - \gamma\right) \Rightarrow 0$$

in $D[t_0, 1]$ for all $t_0 \in (0, 1)$.

2. If $\sqrt{n\overline{F}(u_n)} = o\left(\frac{n}{d_{n,r}}\right)$, then as $n \to \infty$, $\sqrt{n\overline{F}(u_n)t}\left(\widehat{\gamma}_{\lfloor nt \rfloor} - \gamma\right) \Rightarrow \int_1^\infty s^{-1}W\left(s^{-\alpha}, t\right)ds - \alpha^{-1}W\left(1, t\right)$ (2.12) $\sqrt{k_n}t\left(\widehat{\gamma}_{\mathrm{Hill}}(t) - \gamma\right) \Rightarrow \int_1^\infty s^{-1}W\left(s^{-\alpha}, t\right)ds - \alpha^{-1}W\left(1, t\right)$ (2.13)

in $D[t_0, 1]$ for all $t_0 \in (0, 1)$.

- **Remark 3.** 1. The zero limit in the first part of Corollary 1 stems from a degenerate nature of the limiting process in (2.10). Indeed, the limiting process is random in t and deterministic in s.
 - 2. Following Kulik and Soulier (2011), we conjecture that the proper scaling in the first case is $a_n = \sqrt{n\overline{F}(u_n)}$, as well, yielding the same limit as in the second case. However, within the scope of this article, we will not consider the corresponding argument in detail.
 - 3. The limit in (2.12) and (2.13) corresponds to $\gamma B(t), t \in [0, 1]$, where *B* is a standard Brownian motion.

2.4 Asymptotic distribution of the test statistics24

2.4 Asymptotic distribution of the test statistics

Recall that the considered test statistics for the change-point problem (H, A)are defined by

$$\Gamma_n := \sup_{t \in [t_0, 1]} t \left| \frac{\widehat{\gamma}_{\lfloor nt \rfloor}}{\widehat{\gamma}_n} - 1 \right| \text{ and } \widetilde{\Gamma}_n := \sup_{t \in [t_0, 1]} t \left| \frac{\widehat{\gamma}_{\mathrm{Hill}}(t)}{\widehat{\gamma}_{\mathrm{Hill}}(1)} - 1 \right|$$

Using the convergence obtained in Corollary 1 we derive the asymptotic distribution of the test statistics.

Corollary 2. Let X_j , $j \in \mathbb{N}$, be a stationary time series with marginal tail distribution function \overline{F} . Moreover, assume that Assumptions 1 and 2 hold. If $\sqrt{n\overline{F}(u_n)} = o\left(\frac{n}{d_{n,r}}\right)$, then, for all $t_0 \in (0, 1)$, as $n \to \infty$, $\sqrt{n\overline{F}(u_n)} \sup_{t \in [t_0, 1]} t \left| \frac{\widehat{\gamma}_{[nt]}}{\widehat{\gamma}_n} - 1 \right| \Rightarrow \sup_{t \in [t_0, 1]} |B(t) - tB(1)|$, $\sqrt{k_n} \sup_{t \in [t_0, 1]} t \left| \frac{\widehat{\gamma}_{\text{Hill}}(t)}{\widehat{\gamma}_{\text{Hill}}(1)} - 1 \right| \Rightarrow \sup_{t \in [t_0, 1]} |B(t) - tB(1)|$,

where $B(t), t \in [0, 1]$, denotes a standard Brownian motion.

3. Simulations

For all simulations, the following specifications are made:

$$X_j = \sigma(Y_j)\varepsilon_j, \quad j \ge 1 , \qquad (3.14)$$

where

- ε_j, j ≥ 1, is an independent, identically distributed sequence of Pareto distributed random variables generated by the function rgpd (fExtremes package in R);
- Y_j , $j \ge 1$, is a fractional Gaussian noise sequence generated by the function simFGNO (longmemo package in R) with Hurst parameter H;

•
$$\sigma(y) = \exp(y)$$
.

Under the alternative, we insert a change of height h at location $k = \lfloor n\tau \rfloor$ by simulating independent, identically Pareto distributed observations $\varepsilon_j, j \ge 1$, with $\varepsilon_j, j = 1, \ldots, k$, having tail index $\alpha_1 = \ldots = \alpha_k = \alpha$ and $\varepsilon_j, j = k + 1, \ldots, n$, having tail index $\alpha_{k+1} = \ldots = \alpha_n = \alpha + h$. According to Breiman's lemma this induces a change in the tail index of the observations X_1, \ldots, X_n .

We base test decisions on the statistic $\widetilde{\Gamma}_n := \max_{1 \leq k \leq n-1} \Gamma_{k,n}$, where

$$\widetilde{\Gamma}_{k,n} = \frac{k}{n} \left| \frac{\widehat{\gamma}_{\text{Hill}}\left(\frac{k}{n}\right)}{\widehat{\gamma}_{\text{Hill}}(1)} - 1 \right| \text{ with } \widehat{\gamma}_{\text{Hill}}\left(t\right) = \frac{1}{\lfloor k_n t \rfloor} \sum_{i=1}^{\lfloor k_n t \rfloor} \log\left(\frac{X_{\lfloor nt \rfloor : \lfloor nt \rfloor - i + 1}}{X_{\lfloor nt \rfloor : \lfloor nt \rfloor - \lfloor k_n t \rfloor}}\right),$$
(3.15)

and we choose a significance level of 5%.

For the computation of the test statistic, the choice of k_n is considered a delicate issue. In fact, it has been shown in Hall (1982) that the optimal choice of k_n depends on the tail behavior of the data-generating process. Due to this circularity, DuMouchel (1983) suggests to choose k_n proportional to the sample size. As noted in Quintos et al. (2001), a corresponding choice of k_n has been shown to perform well in simulations and is widely used by practitioners. Hence, we choose k_n as p percent of the sample size n, where p = 10% or p = 20%. This is a standard choice in the context of high quantile estimation.

The power of the testing procedures is analyzed by considering different choices for the height of the level shift, denoted by h, and the location of the change-point, denoted by τ . In the tables, the columns that are superscribed by h = 0 correspond to the frequency of a type 1 error, i.e., the rejection rate under the hypothesis.

Both Hurst parameter and tail index, seem to have a significant effect on the rejection rates of the change-point test. An increase in dependence, i.e., an increase of the Hurst parameter H, leads to an increase in the number of rejections. On the one hand, this leads to an increase of power, on the other hand, it results in a larger deviation of the empirical size from the significance level. An increase of tail thickness, i.e., a decrease of the tail parameter α , however, results in an improvement of the test's performance in that the empirical power increases while the empirical size

26

draws closer to the level of significance. Indeed, if the tail is thicker, you have more observations that are informative about the tail, such that tail changes become easier to detect. Moreover, the empirical power of the test is higher for changes to heavier tails, i.e., the test tends to detect changes with a negative change-point height h better.

Considering small values of H and α , i.e., for heavy-tailed time series with unincisive long-range dependence, the simulation results concur with the expected behavior of change-point tests: An increasing sample size goes along with an improvement of the finite sample performance, i.e., the empirical size approaches the level of significance and the empirical power increases; the empirical power of the testing procedures increases when the height of the level shift increases; and the empirical power is higher for breakpoints located in the middle of the sample than for change-point locations that lie close to the boundary of the testing region. A comparison of the rejection rates in Table 1 and 3 reveals that an increase of tail thickness is better detected in the presence of late changes, i.e., when $\tau = 0.75$, than in the presence of early changes, i.e., when $\tau = 0.25$.

27

			$\alpha = 4$								$\alpha = 2$						
	p	n	h = -1	h = -0.5	h = 0	h = 0.5	h = 1	h = -1	h = -0.5	h = 0	h = 0.5	h = 1	h = -1	h = -0.5	h = 0	h = 0.5	h = 1
	0.1	300	10.1	9.4	9.2	8.9	8.8	11.8	9.2	9.1	9.0	8.8	14.5	10.5	9.7	9.9	9.2
9		500	8.8	7.5	7.6	7.0	7.1	9.9	8.3	8.1	7.3	6.9	20.3	10.5	7.8	7.2	7.5
0		1000	6.8	6.6	6.2	5.6	5.3	9.4	6.7	5.7	6.1	5.2	37.3	10.6	6.2	6.1	7.5
Η	0.2	300	7.0	6.5	6.7	6.1	7.1	7.5	7.6	7.3	6.4	6.9	15.6	8.9	7.2	7.4	7.3
		500	6.9	5.9	5.5	5.8	5.7	7.4	6.0	5.5	5.8	5.8	26.2	8.4	6.0	6.2	7.0
		1000	4.8	4.7	4.8	4.6	4.9	7.7	5.1	4.2	4.6	4.6	52.1	9.6	4.7	5.5	7.2
H = 0.7	0.1	300	11.4	11.3	10.5	10.2	11.3	12.2	12.1	10.6	9.8	10.5	16.8	12.5	11.4	10.2	10.4
		500	10.4	9.7	9.7	9.2	9.2	12.5	10.3	9.3	8.8	9.1	20.6	11.0	9.6	9.6	9.1
		1000	9.5	7.9	7.8	7.4	7.6	11.6	8.2	8.1	6.8	7.7	39.3	11.6	7.7	7.6	9.1
	0.2	300	9.1	9.2	9.1	8.9	8.4	10.0	8.9	8.2	8.3	8.6	18.5	10.0	8.0	8.5	9.7
		500	8.1	7.7	8.1	7.4	7.5	8.8	9.0	7.7	7.6	7.8	28.5	10.4	7.3	7.7	8.2
		1000	7.3	7.1	6.8	6.4	6.7	10.3	7.1	6.8	6.6	6.5	57.1	11.4	6.8	6.6	8.8
	0.1	300	15.0	14.6	14.7	14.6	14.2	15.8	14.1	14.3	14.1	12.5	18.6	13.4	13.0	14.1	12.6
œ		500	14.3	14.6	13.6	13.3	13.9	15.4	13.2	13.2	11.4	12.7	21.6	13.5	10.8	12.3	13.7
= 0.		1000	13.6	13.3	13.0	12.6	12.9	16.4	13.5	12.9	11.7	11.5	43.4	14.5	10.7	10.8	12.8
Η	0.2	300	13.9	13.9	13.7	13.2	13.3	13.4	12.5	12.6	12.7	13.0	20.3	12.5	10.7	11.9	10.9
		500	13.8	13.3	12.4	12.3	13.0	14.1	12.6	12.7	12.6	13.7	32.6	13.5	10.1	11.1	12.1
		1000	14.2	12.9	12.9	13.1	13.6	16.9	13.3	12.9	11.4	12.1	62.5	15.7	10.0	11.1	12.9
	0.1	300	19.9	19.3	19.0	19.6	19.3	18.8	17.6	18.9	18.9	18.9	19.2	16.8	16.6	16.5	17.7
6		500	20.6	20.7	20.9	20.6	21.1	19.0	19.1	18.4	18.3	18.6	25.2	16.0	15.2	16.4	17.5
= 0.5		1000	23.1	22.1	22.7	22.2	23.4	24.4	20.2	19.5	19.9	19.2	50.9	20.2	15.6	17.0	20.0
Η	0.2	300	21.0	21.7	20.8	20.0	20.7	19.5	17.7	18.7	18.7	19.3	26.1	16.3	14.0	15.5	17.9
		500	22.7	23.7	22.9	22.8	22.7	22.1	19.6	20.1	19.2	20.2	41.4	18.2	15.3	16.4	19.7
		1000	26.7	26.3	26.2	25.8	27.4	28.7	23.2	22.7	23.7	23.9	75.5	24.3	17.5	20.0	22.2

Table 1: Rejection rates in % of the change-point test based on the statistic $\tilde{\Gamma}_n$, $k_n = \lfloor np \rfloor$, for LMSV time series (Pareto distributed ε_j , $j \ge 1$) of length n with Hurst parameter H, tail index α and a shift in the mean of height h after a proportion $\tau = 0.25$. The calculations are based on 5,000 simulation runs.

			$\alpha = 4$							$\alpha = 3$			$\alpha = 2$					
	p	n	h = -1	h = -0.5	h = 0	h = 0.5	h = 1	h = -1	h = -0.5	h = 0	h = 0.5	h = 1	h = -1	h = -0.5	h = 0	h = 0.5	h = 1	
											-							
	0.1	300	10.5	9.9	10.1	9.3	9.3	13.8	10.3	8.7	8.7	8.5	41.8	14.5	9.7	8.6	8.6	
		500	9.5	8.4	7.5	7.6	7.7	13.5	8.9	7.5	6.7	6.7	60.2	14.8	7.8	8.3	6.9	
= 0.6		1000	8.2	6.4	5.8	5.9	4.9	15.8	8.6	6.6	5.0	5.5	85.3	19.5	6.2	7.2	7.3	
H	0.2	300	8.3	7.4	6.3	7.1	6.4	10.1	8.0	6.7	7.2	6.2	48.5	11.2	7.2	7.3	7.1	
		500	6.8	5.7	6.4	5.4	5.1	10.8	6.6	5.7	5.8	5.5	68.7	12.3	6.0	6.1	7.1	
		1000	6.6	4.7	4.6	4.4	4.2	14.3	5.8	4.4	4.6	4.8	91.1	18.5	4.7	5.3	7.5	
H = 0.7	0.1	300	12.5	11.4	10.9	10.8	10.1	15.3	12.0	10.6	10.0	10.3	44.3	15.9	11.4	10.0	10.2	
		500	11.9	10.5	9.2	9.2	8.5	16.1	11.2	9.4	8.9	8.0	62.6	16.6	9.6	8.2	9.1	
		1000	10.5	8.8	7.4	7.7	7.0	19.1	10.1	7.4	7.3	7.0	86.6	21.3	7.7	7.7	8.6	
	0.2	300	10.2	9.2	8.6	8.2	8.7	13.1	10.0	8.5	8.8	8.3	50.2	13.1	8.0	8.5	9.0	
		500	10.0	8.6	7.1	8.6	7.6	13.7	7.9	7.4	6.8	7.3	70.2	14.7	7.3	7.3	8.1	
		1000	9.2	6.7	7.1	7.2	7.1	17.2	7.4	6.2	6.4	6.4	91.9	21.1	6.8	7.7	9.0	
	0.1	300	15.8	15.2	14.6	12.9	13.0	17.8	14.7	13.2	13.7	13.1	47.7	18.6	13.0	13.5	12.4	
œ		500	16.3	13.7	13.6	13.4	12.6	20.7	15.0	13.0	12.2	11.0	66.2	19.7	10.8	11.9	11.2	
= 0.		1000	16.5	14.6	13.2	12.1	11.4	25.2	14.5	13.3	11.1	10.9	88.1	25.4	10.7	11.3	11.5	
Η	0.2	300	15.0	12.9	12.7	13.5	13.2	17.3	14.3	12.0	12.0	11.0	56.0	17.4	10.7	11.2	11.8	
		500	15.1	13.6	12.5	12.7	12.0	19.7	14.4	12.4	10.5	12.3	74.3	19.2	10.1	10.9	11.8	
		1000	16.8	14.0	13.5	12.7	12.0	25.4	15.2	11.9	11.2	11.5	93.5	26.9	10.0	11.7	14.2	
	0.1	300	22.2	19.6	19.6	20.6	18.3	23.5	19.1	17.7	17.5	17.2	53.0	21.5	16.6	15.1	16.2	
6		500	22.4	21.0	20.1	20.3	18.2	26.4	20.3	17.7	17.3	16.7	71.5	24.4	15.2	15.1	15.9	
= 0.		1000	26.3	22.9	22.8	21.1	21.3	36.0	23.5	19.8	18.4	18.5	92.4	32.2	15.6	16.6	18.5	
Η	0.2	300	22.4	22.3	20.0	19.9	19.9	25.9	19.5	18.3	17.5	18.2	64.6	21.5	14.0	16.2	16.1	
		500	24.3	23.4	21.9	22.0	21.6	29.7	23.0	20.4	20.2	19.8	82.1	27.5	15.3	15.8	18.0	
		1000	30.8	28.0	27.1	25.5	26.5	39.7	27.1	22.7	23.2	22.7	96.6	37.2	17.5	19.2	23.0	

Table 2: Rejection rates in % of the change-point test based on the statistic $\tilde{\Gamma}_n$, $k_n = \lfloor np \rfloor$, for LMSV time series (Pareto distributed ε_j , $j \ge 1$) of length n with Hurst parameter H, tail index α and a shift in the mean of height h after a proportion $\tau = 0.5$. The calculations are based on 5,000 simulation runs.

			$\alpha = 4$							$\alpha = 3$			$\alpha = 2$					
	p	n	h = -1	h = -0.5	h = 0	h = 0.5	h = 1	h = -1	h = -0.5	h = 0	h = 0.5	h = 1	h = -1	h = -0.5	h = 0	h = 0.5	h = 1	
																v		
	0.1	300	10.7	9.4	9.1	9.0	8.8	13.5	9.8	8.4	8.6	8.3	45.8	13.4	9.4	7.7	8.0	
		500	8.5	7.7	6.7	7.2	6.3	12.5	8.7	7.3	7.4	7.1	61.3	14.0	7.6	7.4	7.2	
= 0.6		1000	6.6	6.2	5.6	5.2	5.6	14.6	7.4	6.3	5.4	4.8	83.5	16.5	6.6	5.9	5.8	
H	0.2	300	7.3	7.3	6.2	7.3	6.2	9.6	7.1	6.1	6.0	6.3	45.2	10.6	6.5	6.2	5.9	
		500	6.8	5.8	5.5	5.6	5.2	9.6	6.7	5.8	5.0	5.2	63.3	11.5	6.4	5.1	5.6	
		1000	5.8	5.3	4.6	4.5	3.4	11.3	5.6	4.4	4.7	4.9	87.3	15.0	5.1	5.0	5.8	
H = 0.7	0.1	300	12.7	11.0	12.2	11.4	10.6	15.7	11.5	10.2	10.3	9.9	47.2	15.5	10.0	9.8	9.8	
		500	11.3	9.9	9.3	9.2	9.5	14.8	9.6	10.3	9.6	8.6	60.9	14.7	9.0	8.1	8.7	
		1000	9.7	8.9	8.0	7.6	7.2	16.5	9.8	8.1	6.8	6.9	83.6	18.6	7.6	7.7	7.1	
	0.2	300	10.1	10.5	8.3	8.3	8.4	12.0	9.9	9.1	8.6	8.6	47.2	11.9	8.1	7.9	7.6	
		500	9.5	8.4	7.5	7.8	7.3	12.6	8.1	7.1	7.1	8.1	64.4	14.1	7.8	7.4	6.8	
		1000	8.8	7.0	7.1	6.3	6.2	15.0	8.4	6.3	6.3	6.7	87.4	16.4	6.2	6.2	6.7	
	0.1	300	16.1	14.5	14.9	12.9	14.9	17.9	15.4	13.5	14.0	13.4	49.7	17.4	13.5	12.1	11.7	
œ		500	16.1	14.5	14.1	13.7	12.2	19.6	14.7	13.5	11.9	11.5	64.8	18.9	12.3	10.7	11.0	
= 0.		1000	15.5	14.3	13.3	12.9	12.1	21.8	15.0	11.8	11.4	10.4	84.2	22.3	11.0	10.3	10.6	
Η	0.2	300	15.9	12.8	13.6	13.3	12.9	16.0	13.0	13.0	12.8	11.3	50.3	16.0	10.8	9.9	10.0	
		500	14.7	13.4	12.9	12.3	12.4	19.3	13.6	11.9	11.8	11.7	67.5	17.5	10.1	10.0	9.7	
		1000	15.7	14.1	12.9	12.6	12.1	22.2	13.9	12.4	11.6	11.5	88.2	21.8	9.9	10.0	10.8	
	0.1	300	21.6	20.4	18.6	19.3	19.1	24.4	20.1	17.4	17.4	17.8	55.8	21.5	15.7	14.9	14.5	
6		500	22.4	21.4	20.9	20.0	18.8	25.9	19.5	18.8	17.5	16.7	70.4	23.4	15.3	14.8	14.7	
= 0.		1000	25.5	22.7	21.7	22.1	21.6	31.8	22.0	19.2	18.2	18.7	88.2	29.1	14.5	15.3	15.0	
Η	0.2	300	22.6	21.0	19.1	19.8	20.6	24.0	19.5	18.2	17.8	17.7	61.2	21.7	14.6	13.5	14.2	
		500	24.0	23.1	23.4	22.1	21.2	28.0	21.3	20.2	19.1	18.7	74.7	25.0	15.2	15.4	15.7	
		1000	29.7	26.3	26.9	25.9	27.2	34.9	25.1	22.7	22.7	23.6	91.5	33.5	17.6	16.9	18.9	

Table 3: Rejection rates in % of the change-point test based on the statistic $\tilde{\Gamma}_n$, $k_n = \lfloor np \rfloor$, for LMSV time series (Pareto distributed ε_j , $j \ge 1$) of length n with Hurst parameter H, tail index α and a shift in the mean of height h after a proportion $\tau = 0.75$. The calculations are based on 5,000 simulation runs.

4. Data

The analysis of financial time series, such as stock market prices, usually focuses on log returns instead of the observed data itself. As an example, we consider the log returns of the daily closing indices of Standard & Poor's 500 (S&P 500, in short) defined by

$$L_t := \log\left(\frac{P_t}{P_{t-1}}\right),$$

where P_t denotes the value of the index on day t, in the period from January 2007 to December 2010; see Figure 1. The data set has been obtained from Yahoo Finance and consists of n = 1014 observations.

Comparing the plots of the sample autocorrelation function of the log returns and the sample autocorrelation function of their absolute values in Figure 2, we observe a phenomenon that is often encountered in the context of financial data: the log returns of the index appear to be uncorrelated, whereas the absolute log returns tend to be highly correlated.

Also, the plot in Figure 1 shows that the considered time series exhibits <u>volatility clustering</u>, meaning that large price changes, i.e., log returns with relatively large absolute values, tend to cluster. This indicates that observations are not independent across time, although the absence of linear autocorrelation suggests that the dependence is nonlinear; see Cont (2005).



32

Figure 1: Daily closing index of Standard & Poor's 500 and its log returns from January 2007 to December 2010



33

Figure 2: Sample autocorrelation of the log returns and the absolute log returns of Standard & Poor's 500 daily closing index from January 2007 to December 2010. The two dashed horizontal lines mark the bounds for the pointwise 95% confidence interval of the autocorrelations under the assumption of data generated by white noise.

Another characteristic of financial time series is the occurrence of heavy tails. In particular, probability distributions of log returns often exhibit tails which are heavier than those of a normal distribution. For the S&P 500 data, this property is highlighted by the Q-Q plot in Figure 3.



Figure 3: Q-Q plot for the log returns of Standard & Poor's 500 daily closing index from January 2007 to December 2010.

All of the previously described features of financial data can be covered by the LMSV model considered in our paper.

In view of the fact that the LMSV model captures properties of the log returns of Standard & Poor's 500 daily closing index, we are interested in analyzing the data with respect to a change in the tail index.

We base the test decision on the statistic defined in (3.15). We choose

 $k_n = \lfloor np \rfloor$, i.e., p defines the proportion of the data that the estimation of the tail index is based on. Choosing p = 0.1, the value of the test statistic corresponds to $\tilde{\Gamma}_n = 1.48207$. The 95%-quantile of the limit distribution $\sup_{t \in [0,1]} |B(t) - tB(1)|$ equals 1.3463348. Choosing the critical value for the hypothesis test correspondingly, the value of $\tilde{\Gamma}_n$ therefore indicates a change-point in the tail index at a level of significance of 5%.

A natural estimate for the change-point location is given by that point in time k, where $\Gamma_{k,n}$ attains its maximum. For the considered data, this point in time corresponds to September 16, 2008, i.e., one day after September 15, 2008, the day Lehman Brothers filed for bankruptcy protection; see Figure 4.

Supplementary Materials

The supplementary material contains all the proofs of the paper.

Acknowledgements

Research supported by Collaborative Research Center SFB 823 <u>Statistical</u> modelling of nonlinear dynamic processes.





Figure 4: Log returns of the daily closing index of Standard & Poor's 500 from January 2007 to December 2010. The red dashed line indicates the estimated change-point location.

References

- Beran, J., Feng, Y., Ghosh, S., Kulik, R. (2013). Long-Memory Processes . Springer, Berlin, Heidelberg.
- Betken, A. and Kulik, R. (2019). Testing for change in long-memory stochastic volatility time series. <u>Journal of Time Series Analysis</u>, 40(5):707 – 738.
- Bilayi-Biakana, C., Ivanoff, G., and Kulik, R. (2019). The tail empirical process for long memory stochastic volatility models with leverage. <u>Electronic Journal of Statistics</u>, 13(2):3453 – 3484.
- Breidt, F. J., Crato, N., and de Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. Journal of Econometrics, 83(1-2):325 348.

Breiman, L. (1965). On some limit theorems similar to the arc-sin law. Teoriya Veroyatnostei

i ee Primeneniya, 10:351 – 360.

- Cont, R. (2005). Long range dependence in financial markets. In <u>Fractals in engineering</u>, pages 159 – 179. Springer.
- Deo, R., Hsieh, M., Hurvich, C. M., and Soulier, P. (2006). Long memory in nonlinear processes. In <u>Dependence in probability and statistics</u>, volume 187 of <u>Lecture Notes in Statistics</u>, pages 221 – 244. Springer, New York.
- Drees, H. (1998a). A general class of estimators of the extreme value index. <u>Journal of Statistical</u> Planning and Inference, 66(1):95 – 112.
- Drees, H. (1998b). On smooth statistical tail functionals. <u>Scandinavian Journal of Statistics</u>, 25:187 – 210.
- Drees, H. (1998c). Optimal rates of convergence for estimates of the extreme value index. <u>The</u> Annals of Statistics, 26(1):434 – 448.
- Drees, H. (2000). Weighted approximations of tail processes for β -mixing random variables. Annals of Applied Probability, 10(4):1274 – 1301.
- DuMouchel, W. H. (1983). Estimating the stable index α in order to measure tail thickness: a critique. Annals of Statistics, 11(4):1019 1031.
- Einmahl, J. H. J. (1990). The empirical distribution function as a tail estimator. <u>Statistica</u> Neerlandica, 44(2):79 – 82.

- Einmahl, J. H. J. (1992). Limit theorems for tail processes with application to intermediate quantile estimation. Journal of Statistical Planning and Inference, 32(1):137 145.
- Einmahl, J. H. J., Gantner, M., and Sawitzki, G. (2010). Asymptotics of the shorth plot. Journal of Statistical Planning and Inference, 140(11):3003 – 3012.

Galbraith, J. W. and Zernov, S. (2004). Circuit breakers and the tail index of equity returns.

Journal of Financial Econometrics, 2(1):109 – 129.

- Hall, P. (1982). On some simple estimates of an exponent of regular variation. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 44(1):37 – 42.
- Harvey, A. C. (2002). Long memory in stochastic volatility. In <u>Forecasting Volatility in the</u> Financial Markets, 307 – 320. Butterwoth-Heinemann Finance.
- Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. <u>Annals of Statistics</u>, 3(5):1163 – 1174.
- Hoga, Y. (2017). Change point tests for the tail index of β -mixing random variables. <u>Econometric Theory</u>, 33(4):915 – 954.

Hurvich, C. M. and Soulier, P. (2009). Stochastic Volatility Models with Long Memory, pages 345 – 354. Springer.

Kim, M. and Lee, S. (2011). Change point test for tail index for dependent data. <u>Metrika</u>, 74(3):297 – 311.

Kim, M. and Lee, S. (2012). Change point test of tail index for autoregressive processes. Journal

of the Korean Statistical Society, 41(3):305 - 312.

- Koedijk, K. G., Schafgans, M. M., and De Vries, C. G. (1990). The tail index of exchange rate returns. Journal of International Economics, 29(1-2):93 – 108.
- Kulik, R. and Soulier, P. (2011). The tail empirical process for long memory stochastic volatility sequences. <u>Stochastic Processes and their Applications</u>, 121(1):109 – 134.
- Phillips, P. C., Loretan, M., et al. (1990). Testing covariance stationarity under moment condition failure with an application to common stock returns. Technical report, Cowles Foundation for Research in Economics, Yale University.
- Mandelbrot, B. B. (1963). The Variation of Certain Speculative Prices. <u>Journal of Business</u>, 36:394 – 419.
- Mason, D. M. (1988). A strong invariance theorem for the tail empirical process. In <u>Annales</u> <u>de l'IHP Probabilités et statistiques</u>, volume 24, pages 491 – 506.
- Pipiras, V. and Taqqu, M. S. (2017). <u>Long-Range Dependence and Self-Similarity</u>, volume 45. Cambridge University Press.
- Pollard, D. (1984). <u>Convergence of stochastic processes</u>. Springer Series in Statistics. Springer-Verlag, New York.
- Quintos, C., Fan, Z., and Phillips, P. (2001). Structural change tests in tail behaviour and the Asian crisis. The Review of Economic Studies, 68(3):633 663.

Resnick, S. I. and Stărică, C. (1997). Asymptotic behavior of hill's estimator for autoregressive

data. <u>Communications in Statistics. Stochastic Models</u>, 13(4):703 – 721. Heavy tails and highly volatile phenomena.

Rootzén, H. (2009). Weak convergence of the tail empirical process for dependent sequences. Stochastic Processes and their Applications, 119(2):468 – 490.

Shorack, G. R. and Wellner, J. A. (1986). <u>Empirical processes with applications to statistics</u>. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.

Taqqu, M. S. (1979). Convergence of Integrated Processes of Arbitrary Hermite Rank. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 50(1):53 – 83.

Taylor, S. J. (1986). Modelling Financial Time Series. Wiley, New York.

Werner, T. and Upper, C. (2004). Time variation in the tail behavior of Bund future returns.

Journal of Futures Markets: Futures, Options, and Other Derivative Products, 24(4):387

- 398.

University of Twente

E-mail: a.betken@utwente.nl

Faculty of Mathematics, Ruhr-Universität Bochum

E-mail: davide.giraudo@rub.de

Department of Mathematics and Statistics, University of Ottawa

E-mail: rafal.kulik@uottawa.ca