ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Statistical Methodology
Series B

J. R. Statist. Soc. B (2016)
78, *Part* 1, *pp.* 31–51

Statistics of heteroscedastic extremes

John H. J. Einmahl,

Tilburg University, The Netherlands

Laurens de Haan

Erasmus University Rotterdam, The Netherlands, and University of Lisbon, Portugal

and Chen Zhou

De Nederlandsche Bank, Amsterdam, and Erasmus University Rotterdam, The Netherlands

[Received February 2014. Revised September 2014]

Summary. We extend classical extreme value theory to non-identically distributed observations. When the tails of the distribution are proportional much of extreme value statistics remains valid. The proportionality function for the tails can be estimated non-parametrically along with the (common) extreme value index. For a positive extreme value index, joint asymptotic normality of both estimators is shown; they are asymptotically independent. We also establish asymptotic normality of a forecasted high quantile and develop tests for the proportionality function and for the validity of the model. We show through simulations the good performance of the procedures and also present an application to stock market returns. A main tool is the weak convergence of a weighted sequential tail empirical process.

Keywords: Extreme value statistics; Functional limit theorems; Non-identical distributions; Sequential tail empirical process

1. Introduction

Classical extreme value analysis makes statistical inference on the tail region of a probability distribution, based on independent and identically distributed (IID) observations. Nevertheless, the observed data may violate the assumption of IID observations. Two potential deviations may occur: the observations may exhibit serial dependence or they may be drawn from different distributions. In this paper we consider the latter situation and develop extreme value statistics to handle the case when observations are drawn from different distributions. It will turn out that extreme value statistics go through under mild variation of the underlying distributions and that we can quantify this variation which reflects the frequency of extreme events.

We consider the following model. At 'time' points $i=1,\ldots,n$ we have independent observations $X_1^{(n)},\ldots,X_n^{(n)}$ following various continuous distribution functions $F_{n,1},\ldots,F_{n,n}$ that share a common right end point $x^*=\sup\{x: F_{n,i}(x)<1\}\in(-\infty,\infty]$, and there is a continuous distribution function F with the same right end point and a continuous positive function c defined on $[0,1]$ such that

Address for correspondence: John H. J. Einmahl, Department of Econometrics and Operations Research, Center for Economic Research, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands.
E-mail: j.h.j.einmahl@uvt.nl

© 2015 Royal Statistical Society 1369–7412/16/78031

$$\lim_{x\to x^*}\frac{1-F_{n,i}(x)}{1-F(x)}=c\left(\frac{i}{n}\right), \tag{1.1}$$

uniformly for all $n\in\mathbb{N}$ and all $1\leqslant i\leqslant n$ (see de Haan *et al.* (2014)). We impose the condition

$$\int_0^1 c(s)\,\mathrm{d}s=1. \tag{1.2}$$

This not only makes the functon c uniquely defined but also, similarly to a density, c can now be interpreted as the frequency of extremes. We call this situation *heteroscedastic extremes* analogously to the concept of heteroscedasticity and we call c the '*scedasis function*'. For example, $c\equiv 1$ resembles the uniform or homogeneous density, i.e. we have 'homoscedastic extremes'. Note that the limit relation (1.1) compares only the distribution tails and does not impose any assumption on the central parts of the distributions. It describes a flexible non-parametric model that allows for different scales in the tails, as explained below.

In addition, we assume, as in classical extreme value analysis, that F belongs to the domain of attraction of a generalized extreme value distribution. That means, there is a real number γ and a positive scale function a such that, for all $x>0$,

$$\lim_{t\to\infty}\frac{U(tx)-U(t)}{a(t)}=\frac{x^\gamma-1}{\gamma}, \tag{1.3}$$

where

$$U:=\left(\frac{1}{1-F}\right)^{\leftarrow}$$

and '$\leftarrow$' denotes the left continuous inverse function. The index γ is the extreme value index. Write also

$$U_{n,i}:=\left(\frac{1}{1-F_{n,i}}\right)^{\leftarrow}.$$

Combining the domain of attraction condition with equation (1.1), it can be shown that

$$\lim_{t\to\infty}\frac{U_{n,i}(tx)-U_{n,i}(t)}{a(t)\{c(i/n)\}^\gamma}=\frac{x^\gamma-1}{\gamma}. \tag{1.4}$$

Hence, all $F_{n,i}$ belong to the domain of attraction of the same extreme value distribution. They have the same extreme value index γ but (for $\gamma\neq 0$) different scale functions a, as in limit (1.3), i.e. the effect of the variation in the function c is on the scale of extremes instead of on the extreme value index. If $\gamma=0$ the effect is on the location only.

In what follows we shall restrict ourselves to the heavy-tailed case, i.e. $\gamma>0$. This is done in view of applications in finance. Then, $x^*=\infty$ and the domain of attraction condition (1.3) simplifies to

$$\lim_{t\to\infty}\frac{U(tx)}{U(t)}=x^\gamma. \tag{1.5}$$

Then the analogue of condition (1.4) is

$$\lim_{t\to\infty}\frac{U_{n,i}(tx)}{U(t)\{c(i/n)\}^\gamma}=x^\gamma.$$

In this paper we make the following contributions. First, we propose a non-parametric estimator on the integrated scedasis function $C(s) := \int_0^s c(u)\,\mathrm{d}u$, for $s \in [0, 1]$, and establish its asymptotic behaviour. Moreover, we show that the Hill estimator can still be successfully applied to estimate the extreme value index γ, even though the observations are drawn from different distributions. The joint asymptotic distribution of both estimators is established. The estimators of γ and C are asymptotically independent. Second, we test hypotheses on (the presence of) heteroscedastic extremes. The null hypothesis is $c = c_0$ for some given scedasis function c_0. In particular, rejecting $c \equiv 1$, the null hypothesis of homoscedasticity, confirms that heteroscedasticity of extremes is present. Third, for application purposes, we provide estimators of c and of high quantiles corresponding to $F_{n,i}$. In particular, we obtain the asymptotic normality of the forecasted (i.e. for $i = n + 1$) high quantile. In applications, the evolution in time of the high quantile estimates quantifies the effect of heteroscedasticity on the magnitude of extreme events. All of this is presented in Section 2. In Section 3, we validate our model by testing whether the extreme value index is constant over time. In Section 4 we present a simulation study and apply our results to financial data.

We sketch how we handle heteroscedastic extremes statistically. Consider $X_i^{(n)}$ for $i = 1, \ldots, n$. We impose a high threshold. Then the (local) frequency of the exceedances over the threshold reflects the (local) value of the scedasis function whereas the magnitude of the exceedances reflects the value of the extreme value index.

A crucial tool for developing the asymptotic theory is the sequential tail empirical process (STEP), based on non-identically distributed observations. Similarly to the sequential empirical process (see, for example, section 3.5 in Shorack and Wellner (1986)), the STEP is a bivariate process with one co-ordinate denoting time and the other magnitude. We prove, in Section 5, the weighted convergence of the STEP to a bivariate Wiener process. Since all our estimators and test statistics can be written as functionals of the STEP, their statistical properties follow from this result. The asymptotic theory for the STEP is of independent interest and can be used for analysing other statistical procedures for heteroscedastic extremes. In particular, it can be used for proving asymptotic theory for other extreme value index estimators (even when γ is not positive). Also, other tests for testing heteroscedastic extremes or the constant extreme value index can be analysed by using the STEP. Our test statistics for the constant extreme value index are only the more straightforward candidates.

Our study is comparable with other deviations from the IID assumption in extreme value analysis. In the direction of allowing serial dependence, Leadbetter *et al.* (1983) developed probability theory on extremes of stationary weakly dependent time series. Hsing (1991), Drees (2000) and more recently Drees and Rootzén (2010) further developed statistical tools to handle extreme events for weakly dependent observations. In all these studies, the observations were assumed to be identically distributed. In the direction of allowing heteroscedastic extremes, the early contribution Mejzler (1956) provides a probabilistic theory based on independent, non-identically distributed random variables. With regard to statistical analysis of heteroscedastic extremes, a few studies have explored a trend in the parameters of some *limit* distributions in extreme value theory. Davison and Smith (1990) considered, a linear trend in both shape and scale parameters of generalized Pareto distributions, whereas Coles (2001) dealt with a log-linear trend in the scale parameter of generalized Pareto distributions. No asymptotic analysis of the estimators was provided in these studies. Two other studies have provided estimators on trends in extremes with asymptotic properties. Hall and Tajvidi (2000) estimated non-parametric trends in parameters of generalized Pareto distributions and generalized extreme value distributions by locally parameterizing the trend. They established the asymptotic behaviour of the estimators under locally constant or locally linear trends. Differently, de Haan *et al.* (2014) considered a similar model to that in our study but concentrated on specific parametric trends and required a large number of observations at any time point. Compared with all existing studies on heteroscedastic extremes, our approach differs in one or more of the following three aspects: we deal with an extreme value analysis based on the domain of attraction rather than the limit situation; we do not impose any (local) parametric model on the scedasis function; we provide asymptotic properties of the estimators.

This paper also contributes to the literature on testing whether the extreme value index is constant over time. For example, Quintos *et al.* (2001) investigated whether the extreme value index of financial data is time invariant. The test statistics therein focused only on the tail behaviour of observations. The asymptotic theory of the tests statistics assumes that the observations are IID, which is much more strict than the targeted null hypothesis that the extreme value index is invariant over time. Consequently, the testing procedure there would reject in the case of heteroscedastic extremes where in fact the extreme value index is constant. In contrast, our test considers the much larger heteroscedastic extremes model as the null hypothesis.

2. Estimation and testing within the heteroscedastic extremes model

In this section, we consider statistical inference on the scedasis function c and also estimation of the common extreme value index γ. We begin with estimating the integrated function c, which is defined by $C(s) := \int_0^s c(u)\,\mathrm{d}u$, for $s \in [0, 1]$. Intuitively, by focusing on the observations above a high threshold, the function C should be proportional to the number of exceedances of the threshold in the first $[ns]$ observations. This leads to the following estimator. Order the observations $X_1^{(n)}, \ldots, X_n^{(n)}$ as $X_{n,1} \leqslant \ldots \leqslant X_{n,n}$. For a suitable intermediate sequence $k = k(n)$, i.e.

$$\begin{aligned} \lim_{n\to\infty} k &= \infty, \\ \lim_{n\to\infty} \frac{k}{n} &= 0, \end{aligned} \tag{2.1}$$

we define the estimator

$$\hat{C}(s) := \frac{1}{k} \sum_{i=1}^{[ns]} \mathbf{1}_{\{X_i^{(n)} > X_{n,n-k}\}}. \tag{2.2}$$

When the observations are all different, the estimator can be written in terms of the ranks $R_{n,i} = \Sigma_{j=1}^n \mathbf{1}_{\{X_i^{(n)} \geqslant X_j^{(n)}\}}$, $1 \leqslant i \leqslant n$, as $\hat{C}(s) = (1/k)\Sigma_{i=1}^{[ns]} \mathbf{1}_{\{R_{n,i} > n-k\}}$. Next we define the Hill estimator as usual,

$$\hat{\gamma}_{\mathrm{H}} := \frac{1}{k} \sum_{j=1}^{k} \log(X_{n,n-j+1}) - \log(X_{n,n-k}), \tag{2.3}$$

but note that it is not yet clear that this is a proper estimator of γ in the case of heteroscedastic extremes.

To prove the asymptotic normality of these estimators, we need second-order conditions quantifying the speed of convergence in limits (1.1) and (1.5). Firstly, suppose that there is a positive, eventually decreasing function A_1, with $\lim_{t\to\infty} A_1(t) = 0$, such that, as $x \to \infty$,

$$\sup_{n\in\mathbb{N}} \max_{1\leqslant i\leqslant n} \left| \frac{1 - F_{n,i}(x)}{1 - F(x)} - c\left(\frac{i}{n}\right) \right| = O\left[A_1\left\{\frac{1}{1 - F(x)}\right\}\right]. \tag{2.4}$$

Secondly, suppose that there is a function A_2 and a $\rho<0$ such that, as $t\to\infty$, $A_2(t)$ has either positive or negative sign, $A_2(t)\to 0$, and, for any $x>0$,

$$\lim_{t\to\infty}\frac{U(tx)/U(t)-x^{\gamma}}{A_2(t)}=x^{\gamma}\frac{x^{\rho}-1}{\rho}; \tag{2.5}$$

see de Haan and Stadtmüller (1996). We require, as $n\to\infty$,

$$\begin{aligned}\sqrt{k}\,A_1\{n/(2k)\}&\to 0,\\ \sqrt{k}\,A_2(n/k)&\to 0.\end{aligned} \tag{2.6}$$

(Note that $\sqrt{k}\,A_1(n/k)\to 0$ can be used for all the results, except the asymptotic normality of the estimator of $c(1)$ and its consequences.) We further assume that

$$\lim_{n\to\infty}\sqrt{k}\sup_{|u-v|\leqslant 1/n}|c(u)-c(v)|=0. \tag{2.7}$$

Assumption (2.7) is quite weak: if c is Lipschitz continuous of order at least $\frac{1}{2}$, it is a direct consequence of the fact that $k/n\to 0$, as $n\to\infty$.

The following theorem on the joint asymptotic normality of $\hat{C}$ and $\hat{\gamma}_{\mathrm{H}}$ is our main result.

Theorem 1. Suppose that conditions (1.2), (2.1), (2.4), (2.5), (2.6) and (2.7) hold. Then, under a Skorokhod construction, we have that, as $n\to\infty$,

$$\sup_{0\leqslant s\leqslant 1}|\sqrt{k}\{\hat{C}(s)-C(s)\}-B\{C(s)\}|\to 0 \qquad \text{almost surely}$$

and

$$\sqrt{k}(\hat{\gamma}_{\mathrm{H}}-\gamma)\to\gamma N_0 \qquad \text{almost surely,}$$

with B a standard Brownian bridge and N_0 a standard normal random variable. In addition, B and N_0 are independent.

Remark 1. Observe that the asymptotic variance of the Hill estimator $\hat{\gamma}_{\mathrm{H}}$ does not depend on c; hence it is the same as in the IID data case ($c\equiv 1$). Recall that $\hat{\gamma}_{\mathrm{H}}$ is based on the order statistics and $\hat{C}$ on the ranks. In the IID data case the vector of order statistics and the vector of ranks are independent, yielding the independence of $\hat{\gamma}_{\mathrm{H}}$ and $\hat{C}$. In the case of heteroscedastic extremes we do not have the independence of ranks and order statistics; nevertheless we have asymptotic independence of $\hat{\gamma}_{\mathrm{H}}$ and $\hat{C}$. From the proofs (Section 5 and Appendix A) it follows that the asymptotic independence of $\hat{\gamma}$ and $\hat{C}$ also holds for the other estimators in use for γ (even for the broader case $\gamma\in\mathbb{R}$), i.e. the estimator of the cumulative frequency of extremes and that of the extreme value index are asymptotically independent. In fact, the asymptotic theory for $\hat{C}$ does not require the extreme value condition (1.3).

Next, we present an estimator of the function c by using a kernel estimation method. Let G be a continuous, symmetric kernel function on $[-1,1]$ such that $\int_{-1}^{1}G(s)\,\mathrm{d}s=1$; set $G(s)=0$ for $|s|>1$. Let $h:=h_n>0$ be a bandwidth such that $h\to 0$ and $kh\to\infty$, as $n\to\infty$. Then, the function c can be estimated non-parametrically by

$$\hat{c}(s)=\frac{1}{kh}\sum_{i=1}^{n}\mathbf{1}_{\{X_i^{(n)}>X_{n,n-k}\}}G\left(\frac{s-i/n}{h}\right).$$

This estimator is similar to the usual kernel density estimator.

Instead of estimating c, we can also test the null hypothesis that $c=c_0$ for some given function c_0. This is equivalent to testing $C=C_0$ with $C_0(s):=\int_0^s c_0(u)\,\mathrm{d}u$. An important example is testing the null hypothesis $c \equiv 1$, which corresponds to testing that C is the identity function on $[0, 1]$. By rejecting this null hypothesis, we can conclude that heteroscedasticity of extremes is present. We consider a Kolmogorov–Smirnov-type test statistic

$$T_1 := \sup_{0 \leqslant s \leqslant 1} |\hat{C}(s) - C_0(s)|$$

and a Cramér–von-Mises-type test statistic

$$T_2 := \int_0^1 \{\hat{C}(s) - C_0(s)\}^2 \, \mathrm{d}C_0(s).$$

The following direct corollary to theorem 1 gives the asymptotic distributions of these two test statistics under hypothesis H_0.

Corollary 1. Assume that the conditions of theorem 1 hold with $c = c_0$. Then, as $n \to \infty$,

$$\sqrt{k} T_1 \xrightarrow{\mathrm{d}} \sup_{0 \leqslant s \leqslant 1} |B(s)|,$$
$$k T_2 \xrightarrow{\mathrm{d}} \int_0^1 B^2(s) \, \mathrm{d}s,$$

with B a standard Brownian bridge.

Observe that the limiting random variables have well-known probability distributions that do not depend on c_0. Also, the domain of attraction condition on F does not play a role and thus these tests can be applied to a broader class of probability distributions.

Finally, we present how to estimate high quantiles at a time point i when having heteroscedastic extremes. High quantiles are the quantiles $U_{n,i}(1/p)$ with very small tail probability p. According to limit (1.1), we have

$$p = 1 - F_{n,i}\left\{U_{n,i}\left(\frac{1}{p}\right)\right\} \approx c\left(\frac{i}{n}\right)\left[1 - F\left\{U_{n,i}\left(\frac{1}{p}\right)\right\}\right].$$

Hence we obtain $U_{n,i}(1/p) \approx U\{c(i/n)/p\}$. Therefore, to estimate $U_{n,i}(1/p)$ we combine the estimator of the scedasis function c with the quantile estimator corresponding to the distribution function F (see Weissman (1978)) and obtain

$$\widehat{U_{n,i}\left(\frac{1}{p}\right)} = X_{n,n-k}\left\{\frac{k\,\hat{c}(i/n)}{np}\right\}^{\hat{\gamma}_{\mathrm{H}}}.$$

The high quantile estimator can be extended to *forecasting* tail risks, i.e. we intend to estimate the high quantile of an unobserved random variable in the next period, $X_{n+1}^{(n)}$. Extending the function c continuously in a right neighbourhood of 1 and incorporating time point $i = n+1$ in limit (1.1) lead to the following estimator of the high quantile $U_{n,n+1}(1/p)$ of the unobserved $X_{n+1}^{(n)}$:

$$\widehat{U_{n,n+1}\left(\frac{1}{p}\right)} = X_{n,n-k}\left\{\frac{k\,\hat{c}(1)}{np}\right\}^{\hat{\gamma}_{\mathrm{H}}}.$$

Since the estimator involves $\hat{c}$ at the boundary point 1, we use a boundary kernel as follows:

$$\hat{c}(1) = \frac{1}{kh} \sum_{i=1}^{n} \mathbf{1}_{\{X_i^{(n)} > X_{n,n-k}\}} G_b\left(\frac{1 - i/n}{h}\right),$$

with

$$G_b(x) = \frac{\int_0^1 u^2 G(u)\,\mathrm{d}u - x\int_0^1 u\,G(u)\,\mathrm{d}u}{\frac{1}{2}\int_0^1 u^2 G(u)\,\mathrm{d}u - \left\{\int_0^1 u\,G(u)\,\mathrm{d}u\right\}^2} G(x);$$

see, for example, Jones (1993).

The following theorem yields a subtle asymptotic normality result for the forecasted high quantile. The proof is deferred to Appendix A.

Theorem 2. Let the function c be defined on $[0, 1+\varepsilon]$ for some $\varepsilon > 0$. Suppose that condition (2.4) holds with $i = n+1$ included, and that conditions (1.2), (2.1), (2.5), (2.6) and (2.7) hold. Assume that $|c''(1)| < \infty$ and $\int_{-1}^1 |G''(x)|\,\mathrm{d}x < \infty$, and that $p = p_n$ satisfies $np/k \to 0$, as $n \to \infty$. Then, as $n \to \infty$, with a bandwidth h such that $kh \to \infty$, $hk^{1/5} \to \lambda \in [0, \infty)$, and $\sqrt{h}\,\log\{k/(np)\} \to \beta \in [0, \infty)$, we have that

$$\sqrt{(kh)}\left\{\frac{\widehat{U_{n,n+1}(1/p)}}{U_{n,n+1}(1/p)} - 1\right\} \xrightarrow{\mathrm{d}} N\left[\lambda^{5/2}\frac{\gamma c''(1)}{2c(1)}\int_0^1 x^2 G_b(x)\,\mathrm{d}x, \gamma^2\left\{\frac{\int_0^1 G_b^2(x)\,\mathrm{d}x}{c(1)} + \beta^2\right\}\right].$$

3. Testing whether the extreme value index is constant

Here we consider the validity of our model. In particular we test whether the extreme value index γ is constant over time. The idea is to estimate γ from subsamples and to compare the estimates. Concretely, we write $\hat{\gamma}_{(s_1,s_2]}$ for the Hill estimator based on $X^{(n)}_{[ns_1]+1}, \ldots, X^{(n)}_{[ns_2]}$, for any $0 \leqslant s_1 < s_2 \leqslant 1$. Recall that, when estimating γ from the full sample, we use the highest $k+1$ observations. Correspondingly, the number of high observations that are used in $\hat{\gamma}_{(s_1,s_2]}$ must reflect the heteroscedasticity in extremes. We would like to choose $k^*_{(s_1,s_2]} := k\{C(s_2) - C(s_1)\}$, which is proportional to the frequency of having exceedances in this subsample. In practice, we estimate it with $k_{(s_1,s_2]} := k\{\hat{C}(s_2) - \hat{C}(s_1)\}$. The following theorem shows the joint asymptotic behaviour of these partial Hill estimators. The proof is deferred to Appendix A.

Theorem 3. Assume that the conditions of theorem 1 hold. Then, under a Skorokhod construction, we have that for any $\delta > 0$, as $n \to \infty$,

$$\sup_{0\leqslant s_1<s_2\leqslant 1, s_2-s_1\geqslant\delta}\left|\sqrt{k}(\hat{\gamma}_{(s_1,s_2]} - \gamma) - \gamma\frac{W\{C(s_2)\} - W\{C(s_1)\}}{C(s_2) - C(s_1)}\right| \to 0 \qquad \text{almost surely,}$$

where W is a standard Wiener process on [0, 1]. In addition, the process W and the Brownian bridge B from theorem 1 are independent and $W(1) = N_0$ there.

The first test compares all partial Hill estimators such that $\hat{C}(s_2) - \hat{C}(s_1) \geqslant \delta$, for some given $\delta > 0$, with the estimator using the full sample, $\hat{\gamma}_{(0,1]} = \hat{\gamma}_{\mathrm{H}}$. The test statistic is

$$T_3 := \sup_{0\leqslant s_1<s_2\leqslant 1,\, \hat{C}(s_2)-\hat{C}(s_1)\geqslant\delta}\left|\frac{\hat{\gamma}_{(s_1,s_2]}}{\hat{\gamma}_{\mathrm{H}}} - 1\right|.$$

Alternatively, we consider a test statistic with a limited number of partial Hill estimators. Divide the sample into m blocks, with $m > 1$ fixed. The cut-off points of the blocks are $l_1 \leqslant l_2 \leqslant \ldots \leqslant l_{m-1}$ with $l_j := \sup\{s : \hat{C}(s) \leqslant j/m\}$; set $l_0 = 0$ and $l_m = 1$. We use the partial Hill estimator $\hat{\gamma}_{(l_{j-1},l_j]}$ as above but use the highest $[k/m]+1$ observations in each subsample, since, by construction, $\hat{C}(l_j)-\hat{C}(l_{j-1})$ is approximately $1/m$ for each j. Now define the test statistic as

$$T_4 := \frac{1}{m}\sum_{j=1}^{m}\left(\frac{\hat{\gamma}_{(l_{j-1},l_j]}}{\hat{\gamma}_{\mathrm{H}}}-1\right)^2.$$

Corollary 2. Assume that the conditions of theorem 1 hold. Then, we have that, as $n\to\infty$,

$$\sqrt{k}T_3 \xrightarrow{\mathrm{d}} \sup_{0\leqslant s_1<s_2\leqslant 1,\, s_2-s_1\geqslant\delta}\left|\frac{W(s_2)-W(s_1)}{s_2-s_1}-W(1)\right|,$$

$$kT_4 \xrightarrow{\mathrm{d}} \chi^2_{m-1},$$

with W a standard Wiener process.

The proof is deferred to Appendix A. Observe that the limits do not depend on c or γ.

4. Simulations and application

In this section we shall first examine, through simulations, the finite sample behaviour of the two tests on the scedasis function and also of the forecasted high quantile (Section 4.1). Next, in Section 4.2, we shall apply all the tests to check whether the extreme value index (T_3, T_4) and the scedasis function (T_1, T_2) of a stock market return series are invariant over time and we shall also estimate this function.

4.1. Simulations

We consider four data-generating processes (DGPs) as follows.

(a) DGP 1: observations are IID and follow the standard Fréchet distribution, i.e. $F^{(1)}_{n,i}(x)=\exp(-1/x)$, for $x>0$. Here $c\equiv 1$.
(b) DGP 2: observations are independent, with observation i following a rescaled Fréchet distribution, $F^{(2)}_{n,i}(x)=\exp\{-(0.5+i/n)/x\}$, for $x>0$. Here $c(s)=0.5+s$, for $s\in[0,1]$.
(c) DGP 3: observations are independent, with observation i following a rescaled Fréchet distribution, $F^{(3)}_{n,i}(x)=\exp\{-c(i/n)/x\}$, for $x>0$, with $c(s)=2s+0.5$, for $s\in[0,0.5]$, and $c(s)=-2s+2.5$ for $s\in(0.5,1]$,
(d) DGP 4: observations are independent, with observation i following a rescaled Fréchet distribution, $F^{(4)}_{n,i}(x)=\exp\{-c(i/n)/x\}$, for $x>0$, with $c(s)=0.8$, for $s\in[0,0.4]\cup[0.6,1]$, $c(s)=20s-7.2$ for $s\in(0.4,0.5]$ and $c(s)=-20s+12.8$ for $s\in(0.5,0.6)$.

For each DGP, we simulate 1000 samples of size $n=5000$ and take $k=400$.

We apply the two tests of Section 2 to test whether there are heteroscedastic extremes ($H_0: c\equiv 1$). For each level of significance α (1%, 5% and 10%), we show in Table 1 the total number (out of 1000) of rejections for each DGP. We see that both tests perform well, both under the null hypothesis (DGP 1) and under the alternative (DGPs 2–4). In particular the power is high in most cases. Test 2 performs somewhat better for global deviations from the null hypothesis, whereas test 1 detects the spike alternative a little better.

Next we consider the forecasted high quantile $\widehat{U_{n,n+1}(1/p)}$, for $p=0.02$ and $h=0.1$. We take for G the often used biweight kernel $G(x)=15(1-x^2)^2/16, x\in[-1,1]$. For each DGP we compute $\widehat{U_{n,n+1}(1/p)}/U_{n,n+1}(1/p)-1$ and, using the 1000 samples, find approximations for the bias (the mean) and the variance of this expression; see Table 2. We also report the corresponding asymptotic variance, taking $\beta=0$ in theorem 2, since $\sqrt{h}\,\log(k)\to 0$. We see that the bias is very

Table 1. Number of rejections out of 1000 simulated data sets

DGP	*Numbers of rejections for the following values of* α *and tests:*					
	$\alpha = 1\%$		$\alpha = 5\%$		$\alpha = 10\%$	
	T_1	T_2	T_1	T_2	T_1	T_2
1	8	12	44	47	95	98
2	990	998	998	999	1000	1000
3	455	570	838	921	941	987
4	663	521	930	903	979	978

Table 2. Bias, variance and asymptotic variance for the forecasted high quantile for $p = 0.02$

DGP	*Bias*	*Variance*	*Asymptotic variance*
1	−0.028	0.137	0.128
2	−0.041	0.094	0.085
3	0.023	0.278	0.256
4	0.004	0.167	0.160

small, which agrees with theorem 2 since for all DGPs $c''(1) = 0$ and hence the asymptotic bias is 0. Given the subtle nature of theorem 2, the asymptotic variance approximates the observed variance well. With the statistical difficulty of forecasting high quantiles in mind, we also see that, according to the observed bias and variance, our estimator of $U_{n,n+1}(1/p)$ performs well.

4.2. Application

We apply the estimators proposed and testing procedures to address the question 'Are financial crises nowadays more frequent than before?'. For that purpose, we collect daily loss returns of the Standard and Poor's 500 index from 1988 to 2012 as an indicator for the status of the US financial market over this period. It has been documented in the empirical finance literature that the downside of equity returns follows heavy-tailed distributions; see for example Jansen and de Vries (1991) and Kearns and Pagan (1997). Assuming that the loss returns on each day follow, possibly different, heavy-tailed distributions as in limits (1.1) and (1.5), we test whether the extreme value index of the loss returns is invariant over time. If not rejected, we further test whether the scedasis function is invariant over time.

We start with analysing the full sample from 1988 to 2012, consisting of 6302 observations (2926 days with losses) and use $k = 160$. Tests 3 (with $\delta = \frac{1}{4}$) and 4 (with $m = 4$) both yield p-values that are virtually 0. Hence, we strongly reject the null hypothesis that the extreme value index is invariant over time. We do not need to investigate the scedasis function further as our model is not valid for this data set. (Here and below we choose k heuristically as the midpoint of the first stable part in the plot of the Hill estimator against k.)

The observed structural change in the extreme value index might be attributed to the recent financial crisis. Therefore we continue with a 20-years subsample from 1988 to 2007, consisting of 5043 observations (2348 days with losses). This excludes the recent financial crisis (and the so-called 'black Monday' in 1987) but nevertheless includes other crisis events such as the burst of the Internet bubble at the beginning of the 21st century. We test again the null hypothesis that the extreme value index is invariant during this period by using $k = 130$. Tests 3 and 4 yield p-values 0.98 and 0.76 respectively. Hence, we do not reject the null hypothesis of constant extreme value index. In other words, the magnitude of the crisis, measured by the extreme value index, is not varying during this period.

We further test whether the scedasis function is constant in the subsample from 1988 to 2007. Both test 1 and test 2 report strong evidence rejecting the null (the p-values are virtually 0). Hence, although the magnitude remains at a constant level, the frequency of extremes changes over time. We apply our kernel estimator $\hat{c}$ of Section 2, with again the biweight kernel and

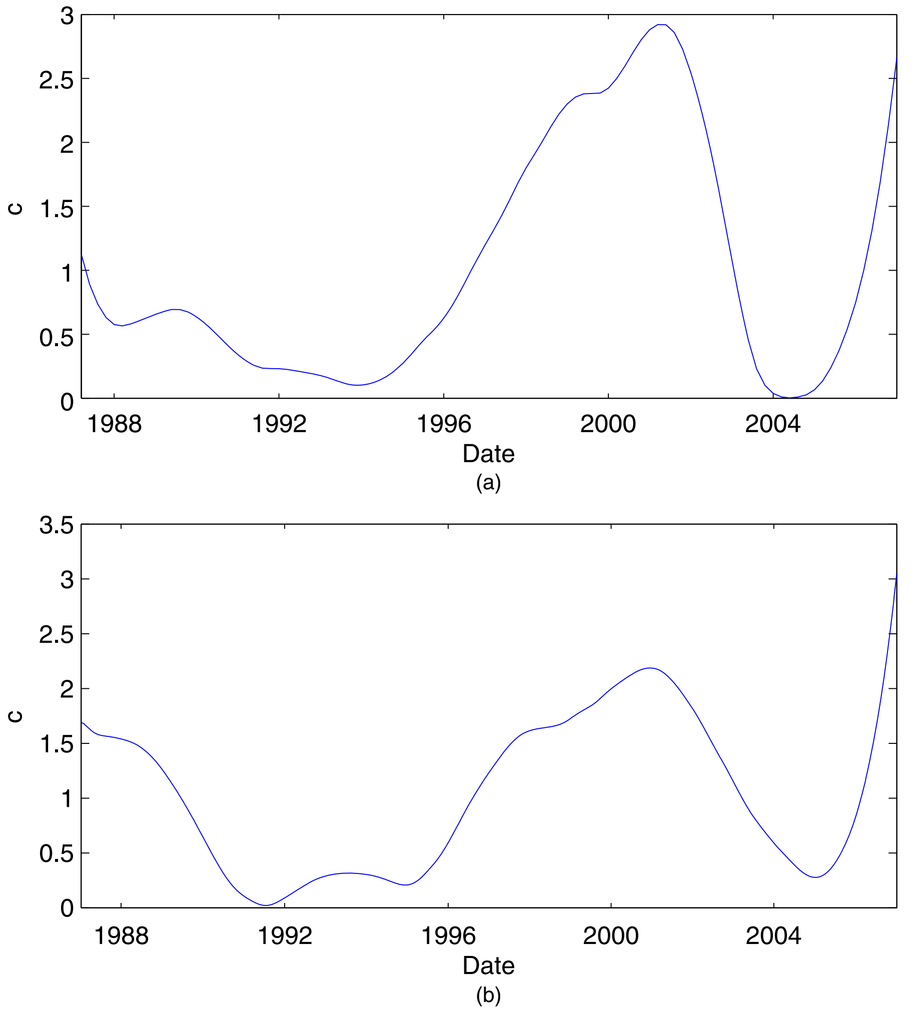

Fig. 1. Estimated scedasis function c based on (a) daily and (b) weekly loss returns of the Standard and Poor's index from 1988 till 2007

$h=0.1$, to estimate the function c. The estimate $\hat{c}$ is plotted in Fig. 1(a). We observe the peak of the scedasis function in the period from 2001 to 2002, which reflects the burst of the Internet bubble. We conclude that the tail risk during these 2 years is higher than that during other periods. Note that, at the end of the period, the scedasis function c increases steeply again, even though we use only data up to the end of 2007, before the financial crisis erupts.

We check the robustness of our results by using *weekly* equity returns. The daily equity return series may suffer from serial dependence such as volatility clustering, which violates our assumption on independence. Such serial dependence is at least much weaker in weekly returns. We repeat our analysis for the weekly loss returns in the subsample from 1988 to 2007, consisting of 1043 observations (454 weeks with losses). Using $k=60$, tests 3 and 4 yield p-values 0.21 and 0.18 respectively. Hence, we do not reject the null hypothesis of constant extreme value index. In addition, tests 1 and 2 yield p-values 0.01 and 0.03 respectively, which provide evidence that the frequency of extremes is time varying during this period. Lastly, with the same kernel estimator $\hat{c}$, we estimate the scedasis function c during this period (Fig. 1(b)). We see from both the quantitative and the qualitative analysis that our results are robust when changing the frequency of the data.

5. The sequential tail empirical process

The proofs of the theorems in Sections 2 and 3 are based on a specific tool: the STEP. In this section, we define the STEP and study its asymptotic properties. Recall that the function c is positive and continuous on $[0,1]$. Thus, there are positive numbers b and d such that $0<b<c(s)<d$, for all $s\in[0,1]$.

Define the sequential empirical distribution function as

$$F_n(x,s):=\frac{1}{n}\sum_{i=1}^{[ns]}\mathbf{1}_{\{X_i^{(n)}\leqslant x\}},\qquad x<x^*.$$

Since we are interested in the right-hand tail of the distribution, we further define the sequential empirical survival function as

$$\bar{F}_n(x,s):=\frac{1}{n}\sum_{i=1}^{[ns]}\mathbf{1}_{\{X_i^{(n)}>x\}}=\frac{[ns]}{n}-F_n(x,s),\qquad x<x^*.$$

Next, we deal with the tail region corresponding to $x=U(n/kt)$, for $0\leqslant t\leqslant 1$, where k satisfies condition (2.1). We approximate the mean and variance of $\bar{F}_n\{U(n/kt),s\}$ as follows. From the limit relation (1.1),

$$\begin{aligned}E\left[\bar{F}_n\left\{U\left(\frac{n}{kt}\right),s\right\}\right]&=\frac{1}{n}\sum_{i=1}^{[ns]}\left[1-F_{n,i}\left\{U\left(\frac{n}{kt}\right)\right\}\right]\\&\approx\frac{1}{n}\sum_{i=1}^{[ns]}c\left(\frac{i}{n}\right)\left[1-F\left\{U\left(\frac{n}{kt}\right)\right\}\right]\approx\frac{kt}{n}C(s).\end{aligned}$$

Similarly, as $n\to\infty$, we obtain the approximation of the variance as

$$\operatorname{var}\left[\bar{F}_n\left\{U\left(\frac{n}{kt}\right),s\right\}\right]=\frac{1}{n^2}\sum_{i=1}^{[ns]}\left[1-F_{n,i}\left\{U\left(\frac{n}{kt}\right)\right\}\right]F_{n,i}\left\{U\left(\frac{n}{kt}\right)\right\}\approx\frac{kt}{n^2}C(s)=O\left(\frac{k}{n^2}\right).$$

Normalizing $\bar{F}_n\,\{U(n/kt),s\}$ with the approximations of its expectation and variance, we define the STEP as

$$\mathbb{F}_n(t,s) := \sqrt{\left(\frac{n^2}{k}\right)}\left[\bar{F}_n\left\{U\left(\frac{n}{kt}\right),s\right\} - \frac{kt}{n}C(s)\right] = \sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{[ns]}\mathbf{1}_{\{X_i^{(n)}>U(n/kt)\}} - t\,C(s)\right\}.$$

We shall prove that, under proper conditions, the STEP converges to a Wiener process in a proper function space.

We start with considering the 'simple' case where F is a standard uniform distribution function and the limit relation (1.1) is exact, i.e., for all $1 \leqslant i \leqslant n$, $1-F_{n,i}(x) = c(i/n)(1-x)$, for $x \in [1-1/c(i/n),1]$. In that case, each $X_i^{(n)}$ follows a uniform distribution on $[1-1/c(i/n),1]$. Hence, we can write $X_i^{(n)} = 1-U_i/c(i/n)$, where the U_i are IID uniform[0,1] random variables. The STEP in this special case is then written as

$$\mathbb{S}_n(t,s) = \sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{[ns]}\mathbf{1}_{\{U_i<c(i/n)kt/n\}} - t\,C(s)\right\}.$$

We call it the *simple* STEP.

We first establish the asymptotic behaviour of the simple STEP. Firstly, we extend the definition of the simple STEP to $(t,s) \in D := (0,2]\times[0,1]$ with the same formula. Secondly, we define a weight function $q(t) = t^\eta$ for any $0 \leqslant \eta < \frac{1}{2}$. Then, we have the following proposition.

Proposition 1. Suppose that k satisfies conditions (2.1) and (2.7). Under a Skorokhod construction, there is a standard bivariate Wiener process $\tilde{W}$ on D, i.e. $\tilde{W}$ is a mean 0 Gaussian process with $\mathrm{cov}\{\tilde{W}(t_1,s_1),\tilde{W}(t_2,s_2)\} = (t_1 \wedge t_2)(s_1 \wedge s_2)$, for $(t_1,s_1),(t_2,s_2) \in D$, such that, as $n \to \infty$,

$$\sup_{(t,s)\in D}\frac{1}{q(t)}|\mathbb{S}_n(t,s) - \tilde{W}\{t,C(s)\}| \to 0 \qquad \text{almost surely.}$$

The proof of this proposition requires the following two lemmas.

Lemma 1. For independent, uniform[0,1] random variables $V_1,\ldots,V_n$, define

$$K_n(t,s) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[ns]}(\mathbf{1}_{\{V_i<t\}} - t), \qquad 0 \leqslant t,s \leqslant 1.$$

Let K denote a Kiefer process on $[0,1]^2$, i.e. K is a mean 0 Gaussian process with

$$\mathrm{cov}\{K(t_1,s_1),K(t_2,s_2)\} = (t_1 \wedge t_2 - t_1t_2)(s_1 \wedge s_2), \qquad \text{for } (t_1,s_1),(t_2,s_2) \in [0,1]^2.$$

Then, we have, under a Skorokhod construction, as $n \to \infty$,

$$\sup_{0<t\leqslant 1,\, 0\leqslant s\leqslant 1}\frac{1}{q(t)}|K_n(t,s) - K(t,s)| \to 0 \qquad \text{almost surely.}$$

Lemma 2. Suppose that $Z_1,\ldots,Z_n$ are independent random variables with Bernoulli distributions: $P(Z_i = 1) = 2c(i/n)k/n$, with k satisfying conditions (2.1) and (2.7). Define the partial sum process as

$$N_n(s) = \sum_{i=1}^{[ns]} Z_i.$$

Then, under a Skorokhod construction, there is a standard Wiener process W_0 on $[0,2]$, such that, as $n \to \infty$,

$$\sup_{0\leqslant s\leqslant 1}\left|\sqrt{k}\left\{\frac{N_n(s)}{k} - 2C(s)\right\} - W_0\{2C(s)\}\right| \to 0 \qquad \text{almost surely.}$$

Lemma 1 follows from theorem 2.12.1 in van der Vaart and Wellner (1996) in combination with the Chibisov–O'Reilly theorem (see page 462 in Shorack and Wellner (1986)). In fact, lemma 2 holds with any non-decreasing continuous function $q:[0,2]\to(0,\infty)$ such that $\int_0^2 u^{-1}\exp\{-\lambda q^2(u)/u\}\,\mathrm{d}u<\infty$, for all $\lambda>0$.

5.1. *Proof of lemma 2*

We apply theorem 2.12.6 in van der Vaart and Wellner (1996) with $Y_{ni}=(1/\sqrt{k})\{Z_i-E(Z_i)\}$, Q_{ni} being equal to the Dirac measure at i/n and Q being equal to a measure on [0, 1] such that $Q([0,s])=2C(s)$. We have that, under a Skorokhod construction, there is a standard Wiener process W_0 on $[0,2]$, such that, as $n\to\infty$,

$$\sup_{0\leqslant s\leqslant 1}\left|\sqrt{k}\left\{\frac{N_n(s)}{k}-2\frac{1}{n}\sum_{i=1}^{[ns]}c\left(\frac{i}{n}\right)\right\}-W_0\{2C(s)\}\right|\to 0 \qquad \text{almost surely.}$$

Lemma 2 is proved provided that

$$\sup_{0\leqslant s\leqslant 1}\sqrt{k}\left|\frac{1}{n}\sum_{i=1}^{[ns]}c\left(\frac{i}{n}\right)-C(s)\right|\to 0 \qquad \text{as } n\to\infty,$$

which follows from assumption (2.7).

5.2. *Proof of proposition 1*

First, we construct n independent uniform[0,1] random variables $U_1,U_2,\ldots,U_n$ in a special way. Recall that d is the upper bound of the function c. For n such that $n/k>2d$, let Z_i, $1\leqslant i\leqslant n$, be independent random variables following Bernoulli distributions with $P(Z_i=1)=2c(i/n)k/n$. Let V_j, $1\leqslant j\leqslant n$, be independent uniform[0,1] random variables, independent of the Z_i. We combine these $2n$ random variables to construct the U_i. Each Z_i is matched with a V_j, where the random index j is defined as follows (recall the notation of lemma 2):

$$j=\begin{cases} N_n(i/n) & \text{if } Z_i=1,\\ i+N_n(1)-N_n(i/n) & \text{if } Z_i=0,\end{cases}$$

i.e. we assign the first $N_n(1)$ random variables V_j to the Z_i with $Z_i=1$, and then assign the rest of the V_j to the Z_i with $Z_i=0$. Then we construct

$$U_i=2Z_i\,c\left(\frac{i}{n}\right)\frac{k}{n}V_j+(1-Z_i)\left[2c\left(\frac{i}{n}\right)\frac{k}{n}+\left\{1-2c\left(\frac{i}{n}\right)\frac{k}{n}\right\}V_j\right], \qquad i=1,\ldots,n.$$

It is straightforward to verify that $U_1,\ldots,U_n$ are independent uniform[0,1] random variables.

We base our simple STEP on these U_i. We then obtain (recalling the notation of lemma 1)

$$\begin{aligned}\mathbb{S}_n(t,s)&=\sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{[ns]}\mathbf{1}_{\{U_i<c(i/n)kt/n\}}-t\,C(s)\right\}\\ &=\sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{N_n(s)}\mathbf{1}_{\{V_i<t/2\}}-t\,C(s)\right\}\\ &=\left\{\frac{N_n(1)}{k}\right\}^{1/2}\frac{1}{\sqrt{N_n(1)}}\sum_{i=1}^{N_n(s)}\left(\mathbf{1}_{\{V_i<t/2\}}-\frac{t}{2}\right)+\frac{t}{2}\sqrt{k}\left\{\frac{N_n(s)}{k}-2C(s)\right\}\end{aligned}$$

$$= \left\{ \frac{N_n(1)}{k} \right\}^{1/2} K_{N_n(1)} \left\{ \frac{t}{2}, \frac{N_n(s)}{N_n(1)} \right\} + \frac{t}{2} \sqrt{k} \left\{ \frac{N_n(s)}{k} - 2C(s) \right\}$$
$$=: I_1(t,s) + I_2(t,s). \tag{5.1}$$

Observe that the two sequences of processes $\{K_m\}$ and $\{N_n\}$ are independent, and hence their limits K and W_0 are independent. We have

$$\frac{1}{q(t)} \left| I_1(t,s) - \sqrt{2} K\left\{ \frac{t}{2}, C(s) \right\} \right| \leqslant \left\{ \frac{N_n(1)}{k} \right\}^{1/2} \frac{1}{q(t)} \left| K_{N_n(1)} \left\{ \frac{t}{2}, \frac{N_n(s)}{N_n(1)} \right\} - K\left\{ \frac{t}{2}, C(s) \right\} \right|$$
$$+ \frac{|K\{t/2, C(s)\}|}{q(t)} \left| \left\{ \frac{N_n(1)}{k} \right\}^{1/2} - \sqrt{2} \right|.$$

Now it readily follows from lemmas 1 and 2 that

$$\sup_{(t,s)\in D} \frac{1}{q(t)} \left| I_1(t,s) - \sqrt{2} K\left\{ \frac{t}{2}, C(s) \right\} \right| \to 0 \qquad \text{almost surely.} \tag{5.2}$$

It is immediate from lemma 2 that, as $n \to \infty$,

$$\sup_{(t,s)\in D} \frac{1}{q(t)} \left| I_2(t,s) - \frac{t}{2} W_0\{2C(s)\} \right| \to 0 \qquad \text{almost surely.} \tag{5.3}$$

Combining results (5.2) and (5.3), yields, as $n \to \infty$,

$$\sup_{(t,s)\in D} \frac{1}{q(t)} \left| \mathbb{S}_n(t,s) - \left[\sqrt{2} K\left\{ \frac{t}{2}, C(s) \right\} + \frac{t}{2} W_0\{2C(s)\} \right] \right| \to 0 \qquad \text{almost surely.}$$

Finally write

$$\tilde{W}(t,s) = \sqrt{2} K\left(\frac{t}{2}, s \right) + \frac{t}{2} W_0(2s),$$

and note that $\tilde{W}$ is a standard bivariate Wiener process on D. □

The following theorem gives the asymptotic behaviour of the STEP in the general case, i.e. in the set-up of Sections 1 and 2.

Theorem 4. Suppose that conditions (1.2), (2.1), (2.4), the first part of condition (2.6) and condition (2.7) hold. Then, under a Skorokhod construction, there is a standard bivariate Wiener process $\tilde{W}$ on $[0,1]^2$ such that, as $n \to \infty$,

$$\sup_{0<t\leqslant 1, 0\leqslant s\leqslant 1} \frac{1}{q(t)} |\mathbb{F}_n(t,s) - \tilde{W}\{t, C(s)\}| \to 0 \qquad \text{almost surely.} \tag{5.4}$$

Proof. Denote $U_i = 1 - F_{n,i}(X_i^{(n)})$. Then $U_1, \ldots, U_n$ are independent, uniform[0,1] random variables. We have, almost surely,

$$\mathbb{F}_n(t,s) = \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{[ns]} \mathbf{1}_{[U_i < 1 - F_{n,i}\{U(n/kt)\}]} - t\,C(s) \right\}.$$

Condition (2.4) implies that there are real numbers $x_0 < x^*$ and $\tau > 0$ such that, for all $x > x_0$, $n \in \mathbb{N}$ and $1 \leqslant i \leqslant n$,

$$c\left(\frac{i}{n}\right) \left[1 - \frac{\tau}{b} A_1\left\{ \frac{1}{1-F(x)} \right\} \right] < \frac{1 - F_{n,i}(x)}{1 - F(x)} < c\left(\frac{i}{n}\right) \left[1 + \frac{\tau}{b} A_1\left\{ \frac{1}{1-F(x)} \right\} \right].$$

Hence,

$$\mathbb{F}_n^-(t,s) \leqslant \mathbb{F}_n(t,s) \leqslant \mathbb{F}_n^+(t,s), \tag{5.5}$$

where

$$\mathbb{F}_n^{\pm}(t,s) := \sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{[ns]} \mathbf{1}_{\{U_i < c(i/n)(kt/n)(1\pm\delta_n)\}} - t\,C(s)\right\},$$

and

$$\delta_n = \sup_{0<t\leqslant 1} \frac{\tau}{b} A_1\left(\frac{n}{kt}\right) = \frac{\tau}{b} A_1\left(\frac{n}{k}\right).$$

Next, we study the asymptotic properties of $\mathbb{F}_n^+$ and $\mathbb{F}_n^-$. With the standard bivariate Wiener process $\tilde{W}$ of proposition 1, we have

$$\begin{aligned}
&\sup_{0<t\leqslant 1,\, 0\leqslant s\leqslant 1} \frac{1}{q(t)} |\mathbb{F}_n^+(t,s) - \tilde{W}\{t, C(s)\}| \\
&\quad\leqslant \sup_{0<t\leqslant 1, 0\leqslant s\leqslant 1} \frac{1}{q(t)} |\mathbb{S}_n^+\{t(1+\delta_n), s\} - \tilde{W}\{t(1+\delta_n), C(s)\}| \\
&\qquad + \sup_{0<t\leqslant 1,\, 0\leqslant s\leqslant 1} \left| \frac{\tilde{W}\{t(1+\delta_n), C(s)\}}{q(t)} - \frac{\tilde{W}\{t, C(s)\}}{q(t)} \right| \\
&\qquad + \sqrt{k}\delta_n \sup_{0<t\leqslant 1,\, 0\leqslant s\leqslant 1} \frac{t}{q(t)} C(s) \\
&\quad =: I_1 + I_2 + I_3.
\end{aligned}$$

From proposition 1 it follows that $I_1 \to 0$ almost surely, as $n \to \infty$. From the (uniform) continuity of the process $\tilde{W}\{t, C(s)\}/q(t)$, extended to $[0,2]\times[0,1]$, we obtain $I_2 \to 0$, as $n \to \infty$. Using $\sqrt{k}\, A_1(n/k) \to 0$ as $n \to \infty$, we obtain $I_3 \to 0$.

Similarly we can show that

$$\sup_{0<t\leqslant 1, 0\leqslant s\leqslant 1} \frac{1}{q(t)} |\mathbb{F}_n^-(t,s) - \tilde{W}\{t, C(s)\}| \to 0 \qquad \text{almost surely.}$$

Now inequality (5.5) yields result (5.4). □

For theorem 4, we did not use the assumption that F belongs to the domain of attraction. With that assumption, we obtain the following corollary.

Corollary 3. Assume that the conditions in theorem 1 hold. Then, for any $0 \leqslant \eta < \frac{1}{2}$ and $x_0 > 0$, under a Skorokhod construction, there is a standard bivariate Wiener process $\tilde{W}$ on $[0, x_0^{-1/\gamma}] \times [0,1]$, such that, as $n \to \infty$,

$$\sup_{0\leqslant s\leqslant 1,\, x\geqslant x_0} x^{\eta/\gamma} \left| \sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{[ns]} \mathbf{1}_{\{X_i^{(n)} > xU(n/k)\}} - x^{-1/\gamma} C(s)\right\} - \tilde{W}\{x^{-1/\gamma}, C(s)\} \right| \to 0 \qquad \text{almost surely.} \tag{5.6}$$

Proof. Set

$$x_n := \frac{n}{k}\left[1 - F\left\{x\, U\left(\frac{n}{k}\right)\right\}\right].$$

By the domain of attraction condition (1.5), we have $x_n \to x^{-1/\gamma}$, as $n \to \infty$, uniformly for all $x \geqslant x_0$. It easily follows from the proof that theorem 4 remains true if we extend the domain of the STEP to $(t,s)\in(0,2x_0^{-1/\gamma}]\times[0,1]$. Therefore, we may replace t in expression (5.4) with x_n to obtain that

$$\sup_{0\leqslant s\leqslant 1,\, x\geqslant x_0} x_n^{-n}\left|\sqrt{k}\left\{\frac{1}{k}\sum_{i=1}^{[ns]}\mathbf{1}_{\{X_i^{(n)}>xU(n/k)\}}-x_n\,C(s)\right\}-\tilde{W}\,\{x_n,C(s)\}\right|\to 0 \qquad \text{almost surely.} \tag{5.7}$$

The proof will be finished once we show that x_n can be replaced by its limit $x^{-1/\gamma}$ at the three places in this expression.

By limit (2.5) we obtain that (see de Haan and Ferreira (2006), page 161) for any $\delta>0$ and sufficiently large n

$$\left|\frac{x_n-x^{-1/\gamma}}{A_2(n/k)}-x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\rho\gamma}\right|\leqslant\delta x^{(-1+\rho)/\gamma}\max(x^{\delta},x^{-\delta}),$$

uniformly for all $x\geqslant x_0$. It follows that

$$\sup_{x\geqslant x_0}\left|\frac{x_n-x^{-1/\gamma}}{A_2(n/k)x^{-1/\gamma}}\right|=O(1), \qquad n\to\infty.$$

Since $A_2(n/k)\to 0$, as $n\to\infty$, we may replace $x_n^{-\eta}$ with $x^{\eta/\gamma}$ in expression (5.7), and since $\sqrt{k}\,A_2(n/k)\to 0$, as $n\to\infty$, we may replace $x_n\,C(s)$ with $x^{-1/\gamma}C(s)$ in expression (5.7). The (uniform) continuity of the weighted bivariate Wiener process implies that, as $n\to\infty$,

$$\sup_{0\leqslant s\leqslant 1,\, x\geqslant x_0} x^{\eta/\gamma}|\tilde{W}\{x_n,C(s)\}-\tilde{W}\{x^{-1/\gamma},C(s)\}|\to 0.$$

Acknowledgements

The research of Laurens de Haan was partially supported by Desenvolvimento de Extremos no Tempo a no Espaço project EXPL/MAT-STA/0622/2013, and by national funds through the Fundação Nacional para a Ciência e Tecnologia, Portugal, under project PEst-OE/MAT/UI0006/2014.

Appendix A: Proofs

A.1. Proof of theorem 1

Taking $s=1$ and $\eta=0$ in expression (5.4) (with domain of t extended to $[0,2]$) yields, as $n\to\infty$,

$$\sup_{0\leqslant t\leqslant 2}\left|\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{n}\mathbf{1}_{\{X_i^{(n)}>U(n/kt)\}}-t\right)-\tilde{W}(t,1)\right|\to 0 \qquad \text{almost surely.}$$

Applying Vervaat's lemma we obtain

$$\sup_{0\leqslant t\leqslant 1}\left|\sqrt{k}\left[\frac{n}{k}\{1-F(X_{n,n-[kt]})\}-t\right]+\tilde{W}(t,1)\right|\to 0 \qquad \text{almost surely.}$$

Taking $t=1$ and denoting $t_n:=(n/k)\{1-F(X_{n,n-k})\}$, we obtain that, as $n\to\infty$,

$$|\sqrt{k}(t_n-1)+\tilde{W}(1,1)|\to 0 \qquad \text{almost surely.} \tag{A.1}$$

We can thus replace t with t_n in expression (5.4) (with domain of t extended to $[0,2]$) and obtain that

$$\sup_{0\leqslant s\leqslant 1}|\sqrt{k}\{\hat{C}(s)-t_n\,C(s)\}-\tilde{W}\{t_n,C(s)\}|\to 0 \qquad \text{almost surely.} \tag{A.2}$$

By applying results (A.1) and (A.2), together with the continuous sample path property of the Wiener process, we obtain that, as $n\to\infty$,

$$\sup_{0\leqslant s\leqslant 1} |\surd k\{\hat{C}(s)-C(s)\}-[\tilde{W}\{1,C(s)\}-C(s)\,\tilde{W}(1,1)]|\to 0 \qquad \text{almost surely.} \tag{A.3}$$

Defining the standard Brownian bridge $B(u)=\tilde{W}(1,u)-u\,\tilde{W}(1,1)$ completes the proof of the first statement in theorem 1.

Next, we prove the second statement: the asymptotic normality of the Hill estimator. Taking $s=1$ and $x_0=\frac{1}{2}$ in expression (5.6) yields, as $n\to\infty$,

$$\sup_{x\geqslant 1/2} x^{\eta/\gamma}\left|\surd k\left(\frac{1}{k}\sum_{i=1}^{n}\mathbf{1}_{\{X_i^{(n)}>xU(n/k)\}}-x^{-1/\gamma}\right)-\tilde{W}(x^{-1/\gamma},1)\right|\to 0 \qquad \text{almost surely.} \tag{A.4}$$

The limit relationship (A.4) is the same as that for the tail empirical process based on IID observations; see de Haan and Ferreira (2006), theorem 5.1.4. Therefore, the asymptotic normality of the Hill estimator, which can be proved via the tail empirical process, follows; see de Haan and Ferreira (2006), example 5.1.5. More precisely, we obtain, as $n\to\infty$, that

$$\surd k(\hat{\gamma}_{\mathrm{H}}-\gamma)\to\gamma\left\{\int_0^1\tilde{W}(t,1)\frac{\mathrm{d}t}{t}-\tilde{W}(1,1)\right\} \qquad \text{almost surely.}$$

It readily follows that $N_0:=\int_0^1\tilde{W}(t,1)\mathrm{d}t/t-\tilde{W}(1,1)$ is standard normal. Finally, it is easy to check that B and $\tilde{W}(\cdot,1)$, and hence B and N_0, are independent. □

Remark 2. It is easy to see that the limit processes in expressions (A.3) and (A.4) are independent. Hence the independence statement of theorem 1 also holds for other estimators of γ based on upper order statistics. In fact similar results can be derived for any real γ.

We proceed with the proof of theorem 3. The proof of theorem 2 will be given at the end of the section.

A.2. Proof of theorem 3

From expression (5.6) we obtain, as $n\to\infty$,

$$\sup_{0\leqslant s_1<s_2\leqslant 1,s_2-s_1\geqslant\delta}\ \sup_{x\geqslant x_0} x^{\eta/\gamma}\left|\surd k\left[\frac{1}{k}\sum_{i=[ns_1]+1}^{[ns_2]}\mathbf{1}_{\{X_i^{(n)}>xU(n/k)\}}-x^{-1/\gamma}\{C(s_2)-C(s_1)\}\right]\right.$$
$$\left.-[\tilde{W}\{x^{-1/\gamma},C(s_2)\}-\tilde{W}\{x^{-1/\gamma},C(s_1)\}]\right|\to 0 \qquad \text{almost surely.} \tag{A.5}$$

From result (A.3), we obtain that, eventually for all s_1 and s_2 such that $s_2-s_1\geqslant\delta$,

$$\hat{C}(s_2)-\hat{C}(s_1)>\tfrac{1}{2}\{C(s_2)-C(s_1)\}>\tfrac{1}{2}b\delta>0 \qquad \text{almost surely.}$$

Hence, dividing expression (A.5) by $\hat{C}(s_2)-\hat{C}(s_1)$, yields, as $n\to\infty$,

$$\sup_{0\leqslant s_1<s_2\leqslant 1,\,s_2-s_1\geqslant\delta}\ \sup_{x\geqslant x_0} x^{\eta/\gamma}\left|\surd k\left\{\frac{1}{k_{(s_1,s_2]}}\sum_{i=[ns_1]+1}^{[ns_2]}\mathbf{1}_{\{X_i^{(n)}>xU(n/k)\}}-x^{-1/\gamma}\frac{C(s_2)-C(s_1)}{\hat{C}(s_2)-\hat{C}(s_1)}\right\}\right.$$
$$\left.-\frac{\tilde{W}\{x^{-1/\gamma},C(s_2)\}-\tilde{W}\{x^{-1/\gamma},C(s_1)\}}{\hat{C}(s_2)-\hat{C}(s_1)}\right|\to 0 \qquad \text{almost surely.} \tag{A.6}$$

Similarly we obtain from result (A.3) that almost surely, as $n\to\infty$,

$$\sup_{0\leqslant s_1<s_2\leqslant 1,\,s_2-s_1\geqslant\delta}\left|\surd k\left\{\frac{\hat{C}(s_2)-\hat{C}(s_1)}{C(s_2)-C(s_1)}-1\right\}-\left[\frac{\tilde{W}\{1,C(s_2)\}-\tilde{W}\{1,C(s_1)\}}{C(s_2)-C(s_1)}-\tilde{W}(1,1)\right]\right|\to 0.$$

Hence, we can replace $\hat{C}(s_2)-\hat{C}(s_1)$ by $C(s_2)-C(s_1)$ in expression (A.6) and obtain that

$$\sup_{0\leqslant s_1<s_2\leqslant 1,\,s_2-s_1\geqslant\delta}\ \sup_{x\geqslant x_0} x^{\eta/\gamma}\left|\surd k\left(\frac{1}{k_{(s_1,s_2]}}\sum_{i=[ns_1]+1}^{[ns_2]}\mathbf{1}_{\{X_i^{(n)}>xU(n/k)\}}-x^{-1/\gamma}\right)-L(x^{-1/\gamma},s_1,s_2)\right|\to 0$$
$$\text{almost surely,} \tag{A.7}$$

where

$$L(v,s_1,s_2):=\frac{\tilde{W}\{v,C(s_2)\}-\tilde{W}\{v,C(s_1)\}}{C(s_2)-C(s_1)}-v\left[\frac{\tilde{W}\{1,C(s_2)\}-\tilde{W}\{1,C(s_1)\}}{C(s_2)-C(s_1)}-\tilde{W}(1,1)\right].$$

Observe that the limit relation (A.7) gives uniformly asymptotic properties of pseudotail empirical processes based on observations from subsamples satisfying $s_2 - s_1 \geqslant \delta$. It is comparable with the limit relation (5.1.18) in de Haan and Ferreira (2006), which is the basis for proving the asymptotic normality of the Hill estimator.

Next, we establish a uniform analogue of the relation (5.1.19) in de Haan and Ferreira (2006). For notational convenience, set $\tilde{k} := k_{(s_1,s_2]}$ and $\tilde{n} := [ns_2] - [ns_1]$. Order the observations $X_{[ns_1]+1}, \ldots, X_{[ns_2]}$ as $X_{s_1,s_2,1} \leqslant \ldots \leqslant X_{s_1,s_2,\tilde{n}}$. We now take $\eta = 0$ in result (A.7) and replace x with $t^{-\gamma}$. Then the generalized Vervaat lemma, as in Einmahl *et al.* (2010), lemma 5, is applied on the collection of functions $(1/k_{(s_1,s_2]}) \Sigma_{i=[ns_1]+1}^{[ns_2]} \mathbf{1}_{\{X_i^{(n)} > t^{-\gamma} U(n/k)\}}$. (Observe that the uniform equicontinuity of the $L(\cdot, s_1, s_2)$ follows from the uniform continuity of W.) In conjunction with the delta method, we obtain

$$\sup_{0 \leqslant s_1 < s_2 \leqslant 1, s_2 - s_1 \geqslant \delta} \; \sup_{1/2 \leqslant t \leqslant 2} \left| \sqrt{k} \left\{ \frac{X_{s_1, s_2, \tilde{n} - [\tilde{k}t]}}{U(n/k)} - t^{-\gamma} \right\} - \gamma t^{-\gamma-1} L(t, s_1, s_2) \right| \to 0 \qquad \text{almost surely,}$$

as $n \to \infty$. Consider $t = 1$: as $n \to \infty$,

$$\sup_{0 \leqslant s_1 < s_2 \leqslant 1, \, s_2 - s_1 \geqslant \delta} \left| \sqrt{k} \left\{ \frac{X_{s_1, s_2, \tilde{n} - \tilde{k}}}{U(n/k)} - 1 \right\} - \gamma L(1, s_1, s_2) \right| \to 0 \qquad \text{almost surely,} \qquad \text{(A.8)}$$

which is a uniform analogue of relation (5.1.19) in de Haan and Ferreira (2006). Using results (A.7) and (A.8) in a similar way to example 5.1.5 in de Haan and Ferreira (2006) yields, as $n \to \infty$,

$$\sup_{0 \leqslant s_1 < s_2 \leqslant 1, s_2 - s_1 > \delta} \left| \sqrt{k} (\hat{\gamma}_{(s_1,s_2]} - \gamma) - \gamma \left\{ \int_0^1 L(u, s_1, s_2) \frac{\mathrm{d}u}{u} - L(1, s_1, s_2) \right\} \right| \to 0 \qquad \text{almost surely.}$$

We have

$$\begin{aligned} &\int_0^1 L(u, s_1, s_2) \frac{\mathrm{d}u}{u} - L(1, s_1, s_2) \\ &\quad = \frac{\int_0^1 \tilde{W}\{u, C(s_2)\} - \tilde{W}\{u, C(s_1)\} \mathrm{d}u/u}{C(s_2) - C(s_1)} - \frac{\tilde{W}\{1, C(s_2)\} - \tilde{W}\{1, C(s_1)\}}{C(s_2) - C(s_1)} \\ &\quad = \frac{\int_0^1 \tilde{W}\{u, C(s_2)\} \mathrm{d}u/u - \tilde{W}\{1, C(s_2)\} - \left[\int_0^1 \tilde{W}\{u, C(s_1)\} \, \mathrm{d}u/u - \tilde{W}\{1, C(s_1)\} \right]}{C(s_2) - C(s_1)}. \end{aligned}$$

The proof is completed by noting that the process W defined by

$$W(s) := \int_0^1 \tilde{W}(u, s) \frac{\mathrm{d}u}{u} - \tilde{W}(1, s)$$

is a standard Wiener process.

A.3. Proof of corollary 2

Combining theorem 1 with theorem 3, we obtain

$$\sup_{0 \leqslant s_1 < s_2 \leqslant 1, \, \hat{C}(s_2) - \hat{C}(s_1) \geqslant \delta} \left| \sqrt{k} \left(\frac{\hat{\gamma}_{(s_1,s_2]}}{\hat{\gamma}_{\mathrm{H}}} - 1 \right) - \left[\frac{W\{C(s_2)\} - W\{C(s_1)\}}{C(s_2) - C(s_1)} - W(1) \right] \right| \to 0 \qquad \text{almost surely.}$$

The asymptotic result for T_3 follows from this in conjunction with again theorem 1 and the continuity of the sample paths of W.

Finally we consider T_4. From theorem 3, theorem 1 and the continuity of the sample paths of W, we obtain

$$\sup_{1 \leqslant j \leqslant m} \left| \sqrt{k} (\hat{\gamma}_{(l_{j-1}, l_j]} - \gamma) - m\gamma \left\{ W\left(\frac{j}{m}\right) - W\left(\frac{j-1}{m}\right) \right\} \right| \to 0 \qquad \text{almost surely,}$$

which implies that

$$\sup_{1\leqslant j\leqslant m}\left|\surd k\left(\frac{\hat\gamma_{(l_{j-1},l_j]}}{\hat\gamma_{\mathrm{H}}}-1\right)-\left[m\left\{W\left(\frac{j}{m}\right)-W\left(\frac{j-1}{m}\right)\right\}-W(1)\right]\right|\to 0 \qquad \text{almost surely.}$$

The asymptotic result for T_4 thus follows. □

For the proof of theorem 2, we need the asymptotic normality of $\hat c(1)$.

Proposition 2. Under the assumptions of theorem 2 we have, as $n\to\infty$,

$$\surd(kh)\,\{\hat c(1)-c(1)\}\xrightarrow{\mathrm{d}} N\left\{\lambda^{5/2}\frac{c''(1)}{2}\int_0^1 x^2\,G_b(x)\,\mathrm{d}x,\,c(1)\int_0^1 G_b^2(x)\,\mathrm{d}x\right\}. \tag{A.9}$$

Remark 3. A similar result for $\hat c(s)$, $s\in(0,1)$, can also be established along the same lines.

A.4. Proof of proposition 2

Recall the notation $t_n=(n/k)\{1-F(X_{n,n-k})\}$. The limit relation (A.1) yields, as $n\to\infty$,

$$\mathbb{P}(k^{1/2}h^{1/4}|t_n-1|\leqslant 1)\to 1. \tag{A.10}$$

Write, for either choice of sign, $t^{\pm}:=1\pm 1/(k^{1/2}h^{1/4})$ and, with G_b^+ and G_b^- denoting the positive and negative parts of G_b respectively,

$$\tilde c^{\pm}(u):=\frac{1}{kh}\sum_{i=1}^n \mathbf{1}_{\{X_i^{(n)}>U(n/ku)\}}\,G_b^{\pm}\left(\frac{1-i/n}{h}\right). \tag{A.11}$$

Then, almost surely, $\hat c(1)=\tilde c^+(t_n)-\tilde c^-(t_n)$. We have, on the event in result (A.10),

$$\tilde c^+(t^-)-\tilde c^-(t^+)\leqslant\tilde c^+(t_n)-\tilde c^-(t_n)\leqslant\tilde c^+(t^+)-\tilde c^-(t^-). \tag{A.12}$$

We shall prove that, as $n\to\infty$,

$$\begin{aligned}\surd(kh)\,\{\tilde c^+(t^+)-\tilde c^-(t^-)-c(1)\}&\xrightarrow{\mathrm{d}} N\left\{\lambda^{5/2}\frac{c''(1)}{2}\int_0^1 x^2\,G_b(x)\,\mathrm{d}x,\,c(1)\int_0^1 G_b^2(x)\,\mathrm{d}x\right\},\\ \surd(kh)\,\{\tilde c^+(t^-)-\tilde c^-(t^+)-c(1)\}&\xrightarrow{\mathrm{d}} N\left\{\lambda^{5/2}\frac{c''(1)}{2}\int_0^1 x^2\,G_b(x)\,\mathrm{d}x,\,c(1)\int_0^1 G_b^2(x)\,\mathrm{d}x\right\}.\end{aligned} \tag{A.13}$$

The proof of result (A.13) requires two results, which follow from routine calculus arguments, using $\int_0^1 G_b(x)\,\mathrm{d}x=1$ and $\int_0^1 x\,G_b(x)\,\mathrm{d}x=0$. As $n\to\infty$,

$$\frac{1}{nh}\sum_{i=1}^n c\left(\frac{i}{n}\right)\,G_b^2\left(\frac{1-i/n}{h}\right)\to c(1)\int_0^1 G_b^2(x)\,\mathrm{d}x. \tag{A.14}$$

and

$$\frac{1}{nh}\sum_{i=1}^n c\left(\frac{i}{n}\right)\,G_b\left(\frac{1-i/n}{h}\right)-c(1)-h^2\frac{c''(1)}{2}\int_0^1 x^2\,G_b(x)\,\mathrm{d}x=O\left(\frac{1}{nh}\right)+o(h^2). \tag{A.15}$$

Now we turn to the proof of the first line of result (A.13). Consider the sum of independent random variables

$$\begin{aligned}I&:=kh\{\tilde c^+(t^+)-\tilde c^-(t^-)\}\\&=\sum_{i=1}^n\left\{\mathbf{1}_{\{X_i^{(n)}>U\{n/(kt^+)\}\}}\,G_b^+\left(\frac{1-i/n}{h}\right)-\mathbf{1}_{\{X_i^{(n)}>U\{n/(kt^-)\}\}}\,G_b^-\left(\frac{1-i/n}{h}\right)\right\}.\end{aligned}$$

Note that, as $n\to\infty$,

$$\operatorname{var}(I)\sim kh\,\frac{1}{nh}\sum_{i=1}^n c\left(\frac{i}{n}\right)\,G_b^2\left(\frac{1-i/n}{h}\right)\sim kh\,c(1)\int_0^1 G_b^2(x)\,\mathrm{d}x.$$

Here the last step follows from result (A.14). Applying the Lindeberg–Feller central limit theorem, we obtain

$$\surd(kh)\left\{\tilde{c}^{+}(t^{+})-\tilde{c}^{-}(t^{-})-\frac{E(I)}{kh}\right\}\xrightarrow{\mathrm{d}} N\left\{0, c(1)\int_0^1 G_b^2(x)\,\mathrm{d}x\right\}. \tag{A.16}$$

Next, a straightforward calculation shows that

$$\begin{aligned}\frac{E(I)}{kh}&=t^{+}\frac{1}{nh}\sum_{i=1}^{n}c\left(\frac{i}{n}\right)G_b^{+}\left(\frac{1-i/n}{h}\right)\left[1+O\left\{A_1\left(\frac{n}{kt^{+}}\right)\right\}\right]\\&\quad-t^{-}\frac{1}{nh}\sum_{i=1}^{n}c\left(\frac{i}{n}\right)G_b^{-}\left(\frac{1-i/n}{h}\right)\left[1+O\left\{A_1\left(\frac{n}{kt^{-}}\right)\right\}\right].\end{aligned}$$

We have $|t^{\pm}-1|=1/(k^{1/2}h^{1/4})=o\{1/\surd(kh)\}$ and expression (2.6) yields

$$\surd(kh)A_1\left(\frac{n}{kt^{\pm}}\right)\leqslant\surd k\,A_1\left(\frac{n}{2k}\right)\rightarrow 0.$$

Combining this with equation (A.15) and using $O(1/nh)+o(h^2)=o\{1/\surd(kh)\}$ yields

$$\frac{E(I)}{kh}-c(1)-h^2\frac{c''(1)}{2}\int_0^1 x^2\,G_b(x)\,\mathrm{d}x=o\left\{\frac{1}{\surd(kh)}\right\}.$$

Hence, since $hk^{1/5}\rightarrow\lambda$, as $n\rightarrow\infty$, we have that

$$\lim_{n\rightarrow\infty}\surd(kh)\left\{\frac{E(I)}{kh}-c(1)\right\}=\lambda^{5/2}\frac{c''(1)}{2}\int_0^1 x^2\,G_b(x)\,\mathrm{d}x.$$

This, in conjunction with result (A.16), yields the first line of expression (A.13). The second line follows similarly.

Combining expression (A.13) with (A.10) and (A.12), yields result (A.9).

A.5. Proof of theorem 2

Write $d_n:=k/np$ and recall that $d_n\rightarrow\infty$, as $n\rightarrow\infty$. We have

$$\begin{aligned}\frac{\widehat{U_{n,n+1}(1/p)}}{U_{n,n+1}(1/p)}&=\left\{\frac{\hat{c}(1)}{c(1)}\right\}^{\gamma}d_n^{\hat{\gamma}_{\mathrm{H}}-\gamma}\frac{X_{n,n-k}}{U(n/k)}\hat{c}(1)^{\hat{\gamma}_{\mathrm{H}}-\gamma}\frac{U(n/k)d_n^{\gamma}}{U(1/p)}\frac{U(1/p)\,c(1)^{\gamma}}{U_{n,n+1}(1/p)}\\&=:I_1I_2I_3I_4I_5I_6.\end{aligned}$$

It can be shown that, as $n\rightarrow\infty$,

$$I_j=1+o_p\{1/\surd(kh)\},\qquad\text{for } j=3,4,$$

and

$$I_j=1+o\{1/\surd(kh)\}\qquad\text{for } j=5,6.$$

The statement for $j=3$ follows from theorem 1 and that for $j=4$ follows from result (A.8). The statements for $j=5$ and $j=6$ follow from conditions (2.4)–(2.7) and the properties of regular variation. We omit further details.

Therefore it suffices to show that $\surd(kh)(I_1I_2-1)$ converges in distribution to the normal distribution specified in theorem 2. From proposition 2 and the delta method, we obtain, as $n\rightarrow\infty$,

$$\surd(kh)(I_1-1)\xrightarrow{\mathrm{d}} N\left\{\lambda^{5/2}\frac{\gamma\,c''(1)}{2\,c(1)}\int_0^1 x^2\,G_b(x)\,\mathrm{d}x,\ \frac{\gamma^2\int_0^1 G_b^2(x)\,\mathrm{d}x}{c(1)}\right\}. \tag{A.17}$$

Also, theorem 1 implies that, as $n \to \infty$,

$$\surd(kh)(I_2 - 1) \xrightarrow{\mathrm{d}} N(0, \beta^2\gamma^2). \tag{A.18}$$

Finally we show that I_1 and I_2 are asymptotically independent. Denote with $\hat{\gamma}^*_{(0,1-h]}$ the Hill estimator based on the first $[n(1-h)]$ observations and with sample fraction $[kC(1-h)]$. Obviously $\hat{\gamma}^*_{(0,1-h]}$ is independent of $\tilde{c}^+(t^-) - \tilde{c}^-(t^+)$ and of $\tilde{c}^+(t^+) - \tilde{c}^-(t^-)$; see result (A.11). Mimicking the proof of theorem 3, we obtain $\hat{\gamma}_{\mathrm{H}} - \hat{\gamma}^*_{(0,1-h]} = o_p(1/\surd k)$ and hence

$$I_2 - d_n^{\hat{\gamma}^*_{(0,1-h]} - \gamma} = o_p\{1/\surd(kh)\}.$$

Using results (A.12) and (A.13), this yields that $\surd(kh)(I_1 I_2 - 1)$ converges in distribution to the convolution of the limits in results (A.17) and (A.18).

References

Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*. New York: Springer.
Davison, A. C. and Smith, R. L. (1990) Models for exceedances over high thresholds (with discussion). *J. R. Statist. Soc.* B, **52**, 393–442.
Drees, H. (2000) Weighted approximations of tail processes for β-mixing random variables. *Ann. Appl. Probab.*, **10**, 1274–1301.
Drees, H. and Rootzén, H. (2010) Limit theorems for empirical processes of cluster functionals. *Ann. Statist.*, **38**, 2145–2186.
Einmahl, J. H. J., Gantner, M. and Sawitzki, G. (2010) Asymptotics of the shorth plot. *J. Statist. Planng Inf.*, **140**, 3003–3012.
de Haan, L. and Ferreira, A. (2006) *Extreme Value Theory: an Introduction*. New York: Springer.
de Haan, L., Klein Tank, A. and Neves, C. (2014) On tail trend detection: modeling relative risk. *Extremes*, to be published.
de Haan, L. and Stadtmüller, U. (1996) Generalized regular variation of second order. *J. Aust. Math. Soc.* A, **61**, 381–395.
Hall, P. and Tajvidi, N. (2000) Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statist. Sci.*, **15**, 153–167.
Hsing, T. (1991) On tail index estimation using dependent data. *Ann. Statist.*, **19**, 1547–1569.
Jansen, D. W. and de Vries, C. G. (1991) On the frequency of large stock returns: putting booms and busts into perspective. *Rev. Econ. Statist.*, **73**, 18–24.
Jones, M. C. (1993) Simple boundary correction for kernel density estimation. *Statist. Comput.*, **3**, 135–146.
Kearns, P. and Pagan, A. (1997) Estimating the density tail index for financial time series. *Rev. Econ. Statist.*, **79**, 171–175.
Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*, vol. 11. New York: Springer.
Mejzler, D. (1956) On the problem of the limit distribution for the maximal term of a variational series. *L'vov Politech. Inst. Naucn. Zap. Ser. Fiz. Mat.*, **38**, 90–109.
Quintos, C., Fan, Z. and Phillips, P. C. B. (2001) Structural change tests in tail behaviour and the Asian crisis. *Rev. Econ. Stud.*, **68**, 633–663.
Shorack, G. R. and Wellner, J. A. (1986) *Empirical Processes with Applications to Statistics*. New York: Wiley.
van der Vaart, A. W. and Wellner, J. A. (1996) *Weak Convergence and Empirical Processes*. New York: Springer.
Weissman, I. (1978) Estimation of parameters and large quantiles based on the k largest observations. *J. Am. Statist. Ass.*, **73**, 812–815.