Trends in Extreme Value Indices

Laurens de Haan & Chen Zhou

Taylor & Francis
Taylor & Francis Group

🔓 OPEN ACCESS | Check for updates

Trends in Extreme Value Indices

Laurens de Haan[a] and Chen Zhou[a,b,c]

[a]Erasmus University Rotterdam, Rotterdam, Netherlands; [b]De Nederlandsche Bank, Amsterdam, Netherlands; [c]Tinbergen Institute, Rotterdam, Netherlands

ABSTRACT

We consider extreme value analysis for independent but nonidentically distributed observations. In particular, the observations do not share the same extreme value index. Assuming continuously changing extreme value indices, we provide a nonparametric estimate for the functional extreme value index. Besides estimating the extreme value index locally, we also provide a global estimator for the trend and its joint asymptotic theory. The asymptotic theory for the global estimator can be used for testing a prespecified parametric trend in the extreme value indices. In particular, it can be applied to test whether the extreme value index remains at a constant level across all observations.

1. Introduction

Extreme value analysis makes statistical inference on the tail region of a distribution function. Balkema and de Haan (1974) show that extreme observations above a high threshold follows approximately a scaled generalized Pareto distribution (GPD). Consequently, one main parameter governs the tail behavior: the shape parameter in the GPD, also known as the extreme value index. Estimation of this parameter is therefore of prime importance for tail inference. Classical extreme value analysis assumes that the observations are independent and identically distributed (iid). Recent studies aim at dealing with the case when observations are drawn from different distributions. In this article, we aim at dealing with non-iid observations: we consider a continuously changing extreme value index and try to estimate the functional extreme value index accurately.

Consider a set of distribution functions $F_s(x)$ for $s \in [0, 1]$ and independent random variables $X_i \sim F_{\frac{i}{n}}(x)$ for $i = 1, \dots, n$. Here $F_s(x)$ is assumed to be continuous with respect to s and x. To perform extreme value analysis, assume that $F_s \in \mathcal{D}_{\gamma(s)}$, where $\mathcal{D}.$ denotes the max-domain of attraction with corresponding extreme value index. In this article, we consider the case that the function γ is positive and continuous on $[0, 1]$. The goal is to estimate the function γ and test the hypothesis that $\gamma = \gamma_0$ for some given function γ_0, based on the observations $X_1, \dots, X_n$.

The idea for estimating $\gamma(s)$ locally is similar to the kernel density estimation. More specifically, we will use only observations X_i in the h-neighborhood of s, that is, $i \in I_n(s) = \{i : \left| \frac{i}{n} - s \right| \le h\}$, where $h := h(n)$ is the bandwidth such that as $n \to \infty$, $h \to 0$ and $nh \to \infty$. Correspondingly there will be $[2nh]$ observations for $s \in [h, 1 - h]$. To apply any known estimators for the extreme value index, such as the Hill

estimator, we choose an intermediate sequence $k := k(n)$ such that as $n \to \infty$, $k \to \infty$ and $k/n \to 0$. Then we use the top $[2kh]$ order statistics among the $[2nh]$ local observations in the h-neighborhood of s to estimate $\gamma(s)$. The local Hill estimator for $\gamma(s)$ is defined as follows. Rank the $[2nh]$ observations X_i with $i \in I_n(s)$ as $X_{1,[2nh]}^{(s)} \le \dots \le X_{[2nh],[2nh]}^{(s)}$. Then,

$$\hat{\gamma}_H(s) := \frac{1}{[2kh]} \sum_{i \in I_n(s)} \left(\log X_i - \log X_{[2nh]-[2kh],[2nh]} \right)^+ . \quad (1)$$

We start with considering the local asymptotic normality. Under some suitable conditions for k and h, we can show that, as $n \to \infty$, for each fixed $s \in (0, 1)$,

$$\sqrt{2kh} \left(\hat{\gamma}_H(s) - \gamma(s) \right) \xrightarrow{d} N(0, (\gamma(s))^2).$$

This result is comparable with the asymptotic normality of the Hill estimator, but now the estimation is based on observations with different extreme value indices. The speed of convergence is $\sqrt{2kh}$ because only the top $[2kh]$ order statistics are used in the estimation.

Next, we consider testing the hypothesis that $\gamma(s) = \gamma_0(s)$ for all $s \in [0, 1]$. Although we are able to estimate the function γ locally, since the local estimators use only local observations, their asymptotic limits are obviously independent. That prevents us from constructing a testing procedure. In addition, the local estimators converges with a slow speed of convergence $1/\sqrt{2kh}$. To achieve the stated goal, we consider the estimation of $\Gamma(s) = \int_0^s \gamma(u) du$ and test the equivalent hypothesis that $\Gamma = \Gamma_0$ where Γ_0 is a function defined as $\Gamma_0(s) = \int_0^s \gamma_0(u) du$ for $s \in [0, 1]$.

The function Γ is estimated by aggregating the local estimators of $\gamma(s)$ to a "global estimator" as follows. Consider

a discretized version of $\hat{\gamma}_H(s)$: $\hat{\gamma}_H\left((2[\frac{s}{2h}]+1)h\right)$. Define the estimator of $\Gamma(s)$ as the integral of the discretized version as follows: for all $0 \leq s \leq 1$,

$$\hat{\Gamma}_H(s) = \int_0^s \hat{\gamma}_H\left((2[\frac{u}{2h}]+1)h\right) du. \qquad (2)$$

Note that $\hat{\gamma}_H(s)$ is only defined for $s \leq 1 - h$. For $u > 1 - h$, we may have $(2[\frac{u}{2h}]+1)h = (2[\frac{1}{2h}]+1)h$. Therefore, we need to extend the range of the estimator $\hat{\gamma}_H(s)$ to accommodate this case: we define $\hat{\gamma}_H((2[1/(2h)]+1)h) := \hat{\gamma}_H((2[1/(2h)]-1)h)$.

Note that $\left\{\hat{\Gamma}_H(s)\right\}_{s \in [0,1]}$ is a stochastic process defined on $[0, 1]$. We shall show that, under the same conditions on k and h used in the local estimation, under a Skorokhod construction, there exists a series of Brownian motions $\{W_n(t)\}$ such that, as $n \to \infty$,

$$\sup_{0 \leq s \leq 1} \left| \sqrt{k}\left(\hat{\Gamma}_H(s) - \Gamma(s)\right) - \int_0^s \gamma(u)dW_n(u) \right| \xrightarrow{P} 0.$$

There are two notable features in this asymptotic relation. Firstly, the convergence is uniformly for all $s \in [0, 1]$. Secondly, the speed of convergence for the estimators $\hat{\Gamma}_H(s)$ is $1/\sqrt{k}$. From these two features, it is possible to construct efficient testing methods to test the null hypothesis that $\Gamma(s) = \Gamma_0(s)$ for all $s \in [0, 1]$ with some given function Γ_0.

Our approach can be regarded as a combination of kernel density estimation and extreme value statistics. To prove the local and global asymptotic normality, we need to combine two limiting procedures as the number of observations tending to infinity. First, the observations used are from a h-neighborhood that is shrinking. Second, within each h-neighborhood, we need to apply a threshold to all observations that is increasing toward infinity. If the h-neighborhood shrinks too fast, there will not be sufficient observations in each neighborhood for statistical inference. If it shrinks too slowly, we would have involved too many observations with vary different extreme value indices such that the estimation is distorted. Therefore, the two limiting procedures have to be balanced such that the resulting estimators possess proper asymptotic behavior.

For that purpose, we assume some conditions regarding the choice of k and h that are related to the speed of variation of the distribution function F_s and the continuity of the extreme value index $\gamma(s)$; see conditions (4)–(7). The first two conditions (4) and (5) are typical assumptions in kernel density estimation and extreme value statistics, respectively. The third condition (6) ensures that the extreme value index function $\gamma(s)$ is sufficiently smooth. In other words, observations in the h-neighborhood have extreme value indices that are not too far off. Notice that $\gamma(s)$ governs the parametric structure of the limit only. In order that a unified "threshold" can be applied to all observations in the h-neighborhood, we further assume the smoothness of the intermediate quantiles as in condition (7). These conditions are not too restrictive, see Example 2.1.

The most close studies to our approach are Gardes and Girard (2010) and Goegebeur, Guillou, and Schorgen (2014). The setups of these two studies are similar to our analysis albeit formulated in a conditional setup. The former focuses on nonstochastic covariates, whereas the latter focuses on random covariates. Both approaches proposed estimators using observations locally and established the local asymptotic normality only. To obtain the local asymptotic normality, the conditions assumed in these two studies are quite different from our conditions and the results obtained therein also differ from our results. Besides, we attempt to establish a global result for $\hat{\Gamma}_H(s)$. The global asymptotic result is necessary for conducting hypothesis testing.

Our article is also related to, but differs from, heteroscedastic extremes. Einmahl, de Haan, and Zhou (2016) model the tail region of distributions of non-iid observations by considering the quotient between tails of different distributions and a common tail. By assuming that such quotients stay positive and finite as one goes further into the tail, the asymptotic constant is called "scedasis." Within such a framework, the extreme value index remains unchanged across the non-iid observations. Compared to the heteroscedastic extremes, we allow for continuously changing extreme value index and try to estimate the functional extreme value index accurately. In our case, the tails of the probability distributions are of different order, that is, the quotient between the tails of the distributions at two locations with different extreme value indices tends to either zero or infinity. In other words, we are dealing with distributions that differ much more than in the heteroscedastic extremes. Therefore, our situation cannot be handled in the same way as in heteroscedastic extremes.

This article is also related to the literature dealing with the variation or trend in extreme value index when considering a purely parametric model such as the generalized extreme value (GEV) distribution or the GPD. First, one may model the trend in the parameters of such models as a specific functional of the covariates; see, for example, Smith (1989) for the GEV model and Davison and Smith (1990) for the GPD model. Second, the trend can also be nonparametrically estimated using various local estimation techniques; see, for example, Davison and Ramesh (2000) using the local likelihood method, Hall and Tajvidi (2000) using the local linearization method, among others. Compared to all these studies, we do not impose a fully parameterized model and therefore maintain a semiparametric approach.

Finally, this article contributes to the literature on testing the null hypothesis of constant extreme value index. Quintos, Fan, and Phillips (2001) and more recently Hoga (2017) considered testing a change point in the tail index. Einmahl, de Haan, and Zhou (2016) proposed two tests for the same purpose. Nevertheless, in all these studies, the main asymptotic result for the constructed tests is under a more restrictive null than having constant extreme value index only. In contrast, we consider a wider null hypothesis potentially including models with constant extreme value index that are excluded from the null of the two existing studies. In addition, our study allows for testing the null hypothesis of having a general prespecified trend in the extreme value index beyond the constant function, such as $\gamma(s) = \gamma_0(s)$ for all s.

We demonstrate the performance of our testing procedure by extensive simulation studies. In addition, our estimation procedure for the $\gamma(s)$ function is also validated when the function differs from a constant function. We apply our developed method to two datasets. The first dataset consist of daily precipitations at Saint-Martin-de-Londres, France from 1976 and

2012. The testing results show that we do not reject the constant extreme value index in this period. The second dataset consist of the losses of the S&P 500 index. The testing results show that we do not reject the constant extreme value index in the period from 1988 to 2012 but do reject this null in a longer period from 1963 to 2012. In the second application, we deal with the presence of serial dependence in the data. Nevertheless, we prefer to postpone the incorporation of serial dependence in the theory in order not to overload the already complicated article.

The article is organized as follows. Our main theorems regarding the local and global estimators are presented in Section 2. The testing procedure is established in Section 3 with simulations. Section 4 is devoted to the applications. Proofs are postponed to the Appendix.

2. Main Theorem

We need the following conditions for obtaining the asymptotic theories of the local and global estimators.

First, we assume the usual second order condition, but uniformly for all $s \in [0, 1]$ as follows. Denote $U_s = (1/(1 - F_s))^{\leftarrow}$ as the quantile function corresponding to the distribution F_s. Suppose there exists a continuous negative function $\rho(s)$ on $[0, 1]$ and a set of auxiliary functions $A_s(t)$ that are continuous with respect to s, such that

$$\lim_{t \to \infty} \frac{\frac{U_s(tx)}{U_s(t)} - x^{\gamma(s)}}{A_s(t)} = x^{\gamma(s)} \frac{x^{\rho(s)} - 1}{\rho(s)}, \quad (3)$$

holds for $x > 1/2$ and uniformly for all $s \in [0, 1]$. A similar uniform second order condition has been adopted in Einmahl and Lin (2006).

Next, we require that the intermediate sequence k and the bandwidth h are properly chosen as follows: there exists some positive constant $\varepsilon > 0$ such that as $n \to \infty$,

$$h = h_n \to 0, k = k_n \to \infty, k_n/n \to 0, \frac{k_n h_n}{|\log h_n|} \to \infty, \quad (4)$$

$$\Delta_{1,n} := \sqrt{k_n} \sup_{0 \leq s \leq 1} \left| A_s \left(\frac{n}{k_n} \right) \right| \to 0, \quad (5)$$

$$\Delta_{2,n} := \sqrt{k_n} \log k_n \sup_{|s_1 - s_2| \leq 2h_n} |\gamma(s_1) - \gamma(s_2)| \to 0, \quad (6)$$

$$\Delta_{3,n} := \sup_{|s_1 - s_2| \leq h_n} \left| \frac{\frac{U_{s_1}\left(\frac{n}{k_n}\right)}{U_{s_2}\left(\frac{n}{k_n}\right)} - 1}{A_{s_2}\left(\frac{n}{k_n}\right)} \right| \to 0, \quad (7)$$

Condition (4) ensures that the number of high order statistics used in each local interval tends to infinity. Condition (5) is the one usually required for extreme value analysis to guarantee to have no asymptotic bias in the estimator. Condition (6) assumes that k_n is compatible with the h_n-variation in the γ function. Condition (7) states that $\left(1 - \frac{k}{n}\right)$-quantiles of distributions are sufficiently smooth in short h-intervals.

The following theorem gives the local asymptotic normality for the estimator $\hat{\gamma}_H(s)$ defined in (1).

Theorem 2.1. Let $X_1, X_2, \ldots, X_n$ be independent random variables with distributions $X_i \sim F_{\frac{i}{n}}(x)$, where $F_s(x)$ is a set of distribution functions defined on $s \in [0, 1]$. Assume that $F_s(x)$ is continuous with respect to s and x and $F_s \in \mathcal{D}_{\gamma(s)}$ where $\gamma(s)$ is a positive continuous function on $[0, 1]$. Assume conditions (3)–(7). Then as $n \to \infty$, we have that for all $s \in (0, 1)$,

$$\sqrt{2kh} \left(\hat{\gamma}_H(s) - \gamma(s) \right) \xrightarrow{d} N(0, (\gamma(s))^2).$$

Remark 2.1. The result of Theorem 2.1 is still valid under weaker conditions if replacing $\sqrt{k}$ in (5) and (6) by $\sqrt{kh}$.

The next theorem gives the asymptotic normality of the global estimator $\hat{\Gamma}_H(s)$ defined in (2).

Theorem 2.2. Assume the same conditions as in Theorem 2.1. Then under a Skorokhod construction, there exists a series of Brownian motions $W_n(s)$ such that as $n \to \infty$,

$$\sup_{s \in [0,1]} \left| \sqrt{k} \left(\hat{\Gamma}_H(s) - \Gamma(s) \right) - \int_0^s \gamma(u) dW_n(u) \right| \xrightarrow{P} 0.$$

We show through an example that the assumptions in (4)–(7) are consistent and not too restrictive.

Example 2.1. Suppose $U_s(t) = t^{\gamma(s)}(1 + t^{\rho(s)})$. Then, we get that

$$\frac{\frac{U_s(tx)}{U_s(t)} - x^{\gamma(s)}}{t^{\rho(s)}} = x^{\gamma(s)} \left(x^{\rho(s)} - 1 \right) \left(1 + t^{\rho(s)} \right)^{-1}.$$

Further, we assume that γ and ρ are continuous function such that $\gamma(s) > 0$, $\rho(s) < 0$ and $\gamma(s) + \rho(s) \neq 0$ for all $s \in [0, 1]$. Denote $\bar{\rho} = \sup_{0 \leq s \leq 1} \rho(s) < 0$, $\underline{\rho} = \inf_{0 \leq s \leq 1} \rho(s) \leq \bar{\rho}$, and assume that $\underline{\rho} > 2\bar{\rho}$. In addition, we assume Lipschitz continuity for the γ and ρ functions: for any $s_1, s_2 \in [0, 1]$, $|\gamma(s_1) - \gamma(s_2)| < c_1 |s_1 - s_2|$ and $|\rho(s_1) - \rho(s_2)| < c_2 |s_1 - s_2|$ for some $c_1, c_2 > 0$. With this setup, the second order condition (3) holds with $A_s(t) = t^{\rho(s)}$.

We find proper choices for $k = k_n$ and $h = h_n$ as follows. First choose $k = k_n = n^\eta$ for some $\eta \in \left(\frac{-\rho}{-\rho+1}, \frac{-2\bar{\rho}}{-2\bar{\rho}+1} \right)$. Here we use the assumption $\underline{\rho} > 2\bar{\rho}$. Next, we choose $h = h_n = n^{-\xi}$ where $\xi > 0$ is determined as follows. Since $\eta > \frac{-\rho}{-\rho+1}$ implies that $\eta > (-\rho)(1 - \eta)$, we can choose a positive value ξ such that $\eta > \xi > \max((-\underline{\rho})(1 - \eta), \eta/2)$. We now verify that with these choices, conditions (4)–(7) holds.

First, condition (4) holds due to the fact that $\eta > \xi$.

Second, condition (5) holds if $\sqrt{k} \left(\frac{n}{k} \right)^{\bar{\rho}} \to 0$ as $n \to \infty$. This is guaranteed by the fact that $\eta < \frac{-2\bar{\rho}}{-2\bar{\rho}+1}$.

Third, condition (6) holds if $h\sqrt{k} \log k \to 0$ as $n \to \infty$. This is guaranteed by the fact that $\xi > \eta/2$.

Finally, we find necessary conditions to ensure condition (7). Note that

$$\frac{\frac{U_{s_1}\left(\frac{n}{k}\right)}{U_{s_2}\left(\frac{n}{k}\right)} - 1}{A_{s_2}\left(\frac{n}{k}\right)} = \left[\frac{\left(\frac{n}{k}\right)^{\gamma(s_1) - \gamma(s_2)} - 1}{\left(\frac{n}{k}\right)^{\rho(s_2)}} + \left(\frac{n}{k}\right)^{\gamma(s_1) - \gamma(s_2) + \rho(s_1) - \rho(s_2)} \right.$$

$$\left. - 1 \right] \left(1 + \left(\frac{n}{k}\right)^{\rho(s_2)}\right)^{-1}.$$

Therefore, (7) holds if as $n \to \infty$,

$$h\left(\log\left(\frac{n}{k}\right)\right)\left(\frac{n}{k}\right)^{-\rho} \to 0.$$

This is guaranteed by the fact that $\xi > (-\underline{\rho})(1 - \eta)$.

To conclude, we have shown that our required conditions are consistent. Notice that this example can be easily generalized to $U_s(t) = C(s)t^{\gamma(s)}(1 + D(s)t^{\rho(s)})$, with some proper continuous function $C(s) > 0$ and $D(s)$ on $[0, 1]$. Further the power in the Lipschitz continuity condition of γ and ρ could be any positive number, not necessarily one. Therefore, our required conditions are not too restrictive.

3. Testing Trends in Extreme Value Indices

Theorem 2.2 provides the possibility to test if the extreme value indices follow a specific trend, that is, $H_0 : \gamma(s) = \gamma_0(s)$ for all $s \in [0, 1]$, with some given function γ_0. Similar to testing the specific trend in the "skedasis" function in Einmahl, de Haan, and Zhou (2016), we apply an equivalent test to test $H_0 : \Gamma(s) = \Gamma_0(s) \in [0, 1]$ for all s, where $\Gamma_0(s) = \int_0^s \gamma_0(u)du$. Clearly, one may construct a Kolmogorov–Smirnov (KS) type test with the testing statistic defined as

$$T := \sup_{s \in [0,1]} \left|\hat{\Gamma}_H(s) - \Gamma_0(s)\right|.$$

Then, Theorem 2.2 implies that under the null hypothesis $H_0 : \Gamma(s) = \Gamma_0(s)$ for all $s \in [0, 1]$, as $n \to \infty$,

$$\sqrt{k}T \overset{d}{\to} \sup_{s \in [0,1]} \left|\int_0^s \gamma(u)dW(u)\right|,$$

where $W(u)$ is a standard Brownian motion defined on $[0,1]$.

It is often of interest to test whether the extreme value index remains constant over time, without prior knowledge on the constant extreme value index, that is, $H_0 : \gamma(s) = \gamma$ for all $s \in [0, 1]$ without specifying γ. In this case, one may use $\hat{\Gamma}_H(1)$ as an estimator of the constant extreme value index γ and define the testing statistic as

$$\tilde{T} := \sup_{s \in [0,1]} \left|\frac{\hat{\Gamma}_H(s)}{\hat{\Gamma}_H(1)} - s\right|.$$

It is straightforward to show that under the null hypothesis $H_0 : \gamma(s)$ is a constant for all $s \in [0, 1]$, as $n \to \infty$,

$$\sqrt{k}\tilde{T} \overset{d}{\to} \sup_{s \in [0,1]} |B(s)|,$$

where $B(s)$ is a standard Brownian bridge defined on $[0,1]$. Note that the limit distribution is identical to that in the classical KS test.

We run a simulation study to demonstrate the finite sample performance of the testing procedure using $\tilde{T}$. In all our simulations we generate $m = 2000$ samples with n observations in each sample. We start with setting the sample size to $n = 5000$. For the two parameters k and h, we choose several combinations between $k = 100, 200$ and $h = 0.025, 0.04$.

For each sample, we simulate the observations from the following data generating process

$$X_i = Z_i^{1/\gamma(i/n)}, \quad i = 1, 2, \ldots, n,$$

where $\{Z_i\}_{i=1}^n$ are iid observations from the standard Fréchet distribution with distribution function $F(x) = \exp(-1/x)$ for $x > 0$. For the function $\gamma(s)$ we consider either a linear trend as $\gamma(s) = 1 + bs$ or a trend following the sin function as $\gamma(s) = 1 + c\sin(2\pi s)$. If $b = 0$ or $c = 0$, the two model resemble the iid case, that is, the null hypothesis that the extreme value indices remain constant holds. We consider four alternative cases: $b = 1$, $b = 2$, $c = 1/4$, and $c = 1/2$. In total, we have 20 sets of simulations due to the various choices of k, h and the model of $\gamma(s)$.

For each simulated sample j we apply the test $\tilde{T}$ in Section 3 to test whether the extreme value indices remain constant and obtain the corresponding p-value, p_j, for $j = 1, 2, \ldots, m$. For the simulations based on $b = 0$ (or $c = 0$), that is, when the null hypothesis holds, we make QQ-plots between the simulated p-values across all m samples that are below 0.1 against a uniform distribution on $[0,0.1]$. If the size of the test is agrees with the significance level, the dots in the QQ-plots should line up on the 45 degree line. Figure 1 presents four QQ-plots corresponding to four choices of (k, h). The plots confirm the validity of our test under the null hypothesis.

Next, for all sets of simulations, we calculate the rejection rate based on each significance level α as $\#\{j : p_j < \alpha\}/m$ for $\alpha = 0.01, 0.05$, and 0.1. The rejection rates are reported in Table 1.

In the first panel, we observe that under the null hypothesis, the rejection rates, that is, the Type I error, are close to the significance levels. The difference between the two choices of h is negligible when choosing $k = 200$. The difference between the two choices of k is also negligible when considering $h = 0.025$. For $k = 100$ and $h = 0.04$, the test is conservative.

In the next two panels, the rejections rates can be read as the power of the test. Between the two choices of h, $h = 0.025$ leads to a slightly higher power for rejecting the linear trend, while $h = 0.04$ leads to a slightly higher power for the sin trend. Between the two choices of k, $k = 200$ leads to a much higher power in all alternative models. Therefore, choosing a higher k is preferred as long as the bias is not an issue, whereas the choice of h depends on the shape of the trend.

When comparing across the models, the power is higher for $b = 2$ ($c = 1/2$) than for $b = 1$ ($c = 1/4$). This is in line with the intuition that the test is more powerful to detect larger deviation from the null hypothesis of having constant extreme value index.

Finally, for the two sin trends: $c = 1/4$ and $c = 1/2$, we plot the average of the estimated $\gamma(s)$ across the m samples (the solid line) and its corresponding 95% confidence interval for each given s. There are two ways to construct the confidence interval. First, we use the asymptotic theory in Theorem 2.1 to construct the confidence interval based on the averaged estimate of $\gamma(s)$ (the dotted lines). Second, we can obtain an empirical confidence interval from the m estimates (the dash-dotted lines). The comparison between the two provides a validation of our asymptotic theory. In this exercise, we fix $k = 200$ and $h = 0.025$.

Figure 2 shows the estimation result for the sin trends. Firstly, the average estimates across m sample resembles the true value of the $\gamma(s)$ function (the dash line). Secondly, the
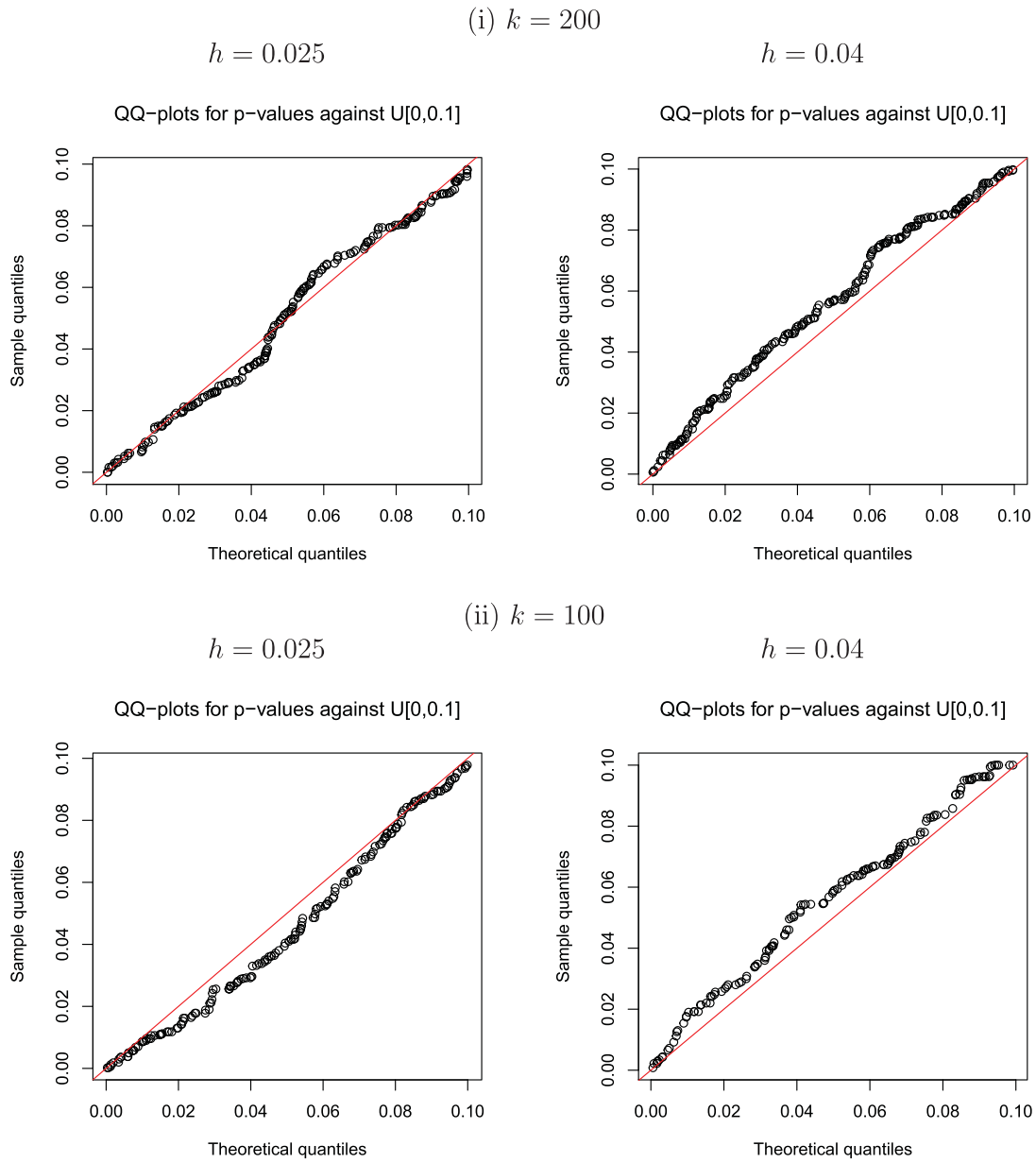
(i) $k = 200$

$h = 0.025$

$h = 0.04$

(ii) $k = 100$

$h = 0.025$

$h = 0.04$

Figure 1. QQ-plots for simulated p-values below 0.1

Table 1. Rejection rates in simulations: sample size $n = 5000$.

|			$k = 200$		$k = 100$	
		α	$h = 0.025$	$h = 0.040$	$h = 0.025$	$h = 0.040$
iid observations		0.1	0.104	0.117	0.092	0.071
		0.05	0.051	0.053	0.052	0.029
		0.01	0.011	0.010	0.013	0.006
Linear trend	$b = 1$	0.1	0.831	0.682	0.539	0.375
		0.05	0.731	0.559	0.407	0.262
		0.01	0.505	0.338	0.207	0.095
	$b = 2$	0.1	0.989	0.960	0.843	0.703
		0.05	0.970	0.917	0.743	0.584
		0.01	0.888	0.740	0.480	0.271
Sin trend	$c = 1/4$	0.1	0.500	0.696	0.254	0.393
		0.05	0.388	0.597	0.165	0.292
		0.01	0.195	0.385	0.064	0.126
	$c = 1/2$	0.1	0.991	0.999	0.850	0.932
		0.05	0.976	0.994	0.770	0.886
		0.01	0.921	0.984	0.533	0.687

(i) $c = 1/4$

(ii) $c = 1/2$
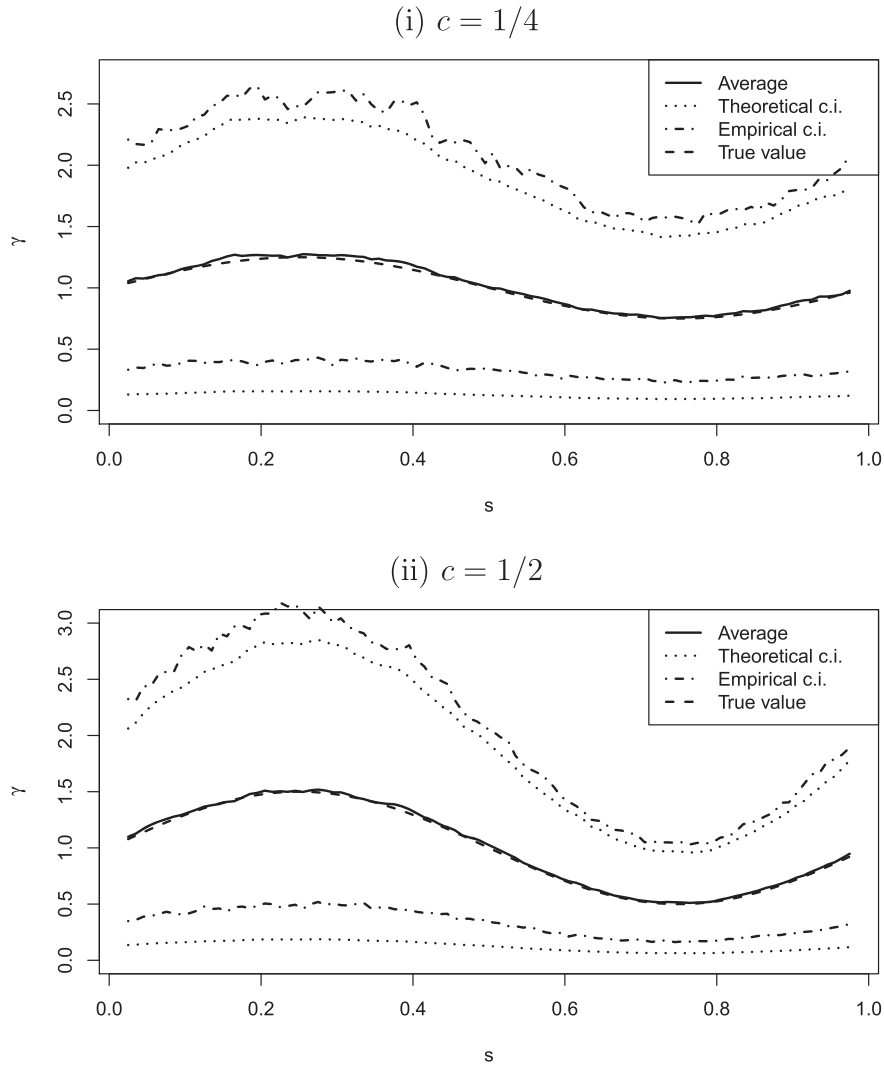
Figure 2. Estimated varying extreme value indices: sin trends.

confidence intervals derived from our asymptotic theory are close to that obtained from the simulation indicating the validity of our asymptotic theory. In both cases, the empirical confidence interval is shifted slightly upward compared to the theoretical confidence interval. This shift cannot be explained by the estimation bias because the average estimates is close to the true value. An alternative explanation is that the asymptotic normality requires large $2kh$, but in this simulation, $2kh = 10$ is rather low. To validate this reasoning, we make QQ-plots for the estimates of $\gamma(1/2)$ against the standard normal distribution in Figure 3. The solid line has a slope of $1/\sqrt{2kh}$. The deviation of the dots from the solid line indicates that the current level of $2kh$ is rather low for having normality of the estimates. Although the middle part of the distribution might be close to a normal distribution, both left and right tails are deviating from normality.

Besides working with large sample size $n = 5000$, we also consider a smaller sample size: $n = 2000$. For the small sample size, we keep $h = 0.025$ and $h = 0.04$, but choose different levels of k at $k = 100$ and $k = 50$. The results for rejection rates is reported in Table 2. The general patterns observed in the simulations for $n = 5000$ are preserved.

4. Application

We apply our developed method to two datasets to test whether the extreme value indices remain unchanged over time. If not rejecting the null, we estimate the constant extreme value index. If rejecting the null, we estimate the time variation in the extreme value indices. Throughout the application analysis, we choose $h = 0.025$.

4.1. Application 1: Precipitation at Saint-Martin-de-Londres

We employ a dataset consisting of the precipitation at Saint-Martin-de-Londres from 1976 to 2015, with 14,610 daily observations. We test the constancy of the extreme value indices over the entire period. The obtained p-values against various levels of k are shown in the *upper panel* of Figure 4. We do not reject the null hypothesis under the 5% significance level (the dash line). We then estimate the constant extreme value index by applying the Hill estimator to all observations, that is, estimating $\Gamma(1)$. The obtained estimates against various levels of k are shown in the *lower panel* of Figure 4. By choosing $k = 200$, we get the estimate for the constant extreme value index at 0.395.
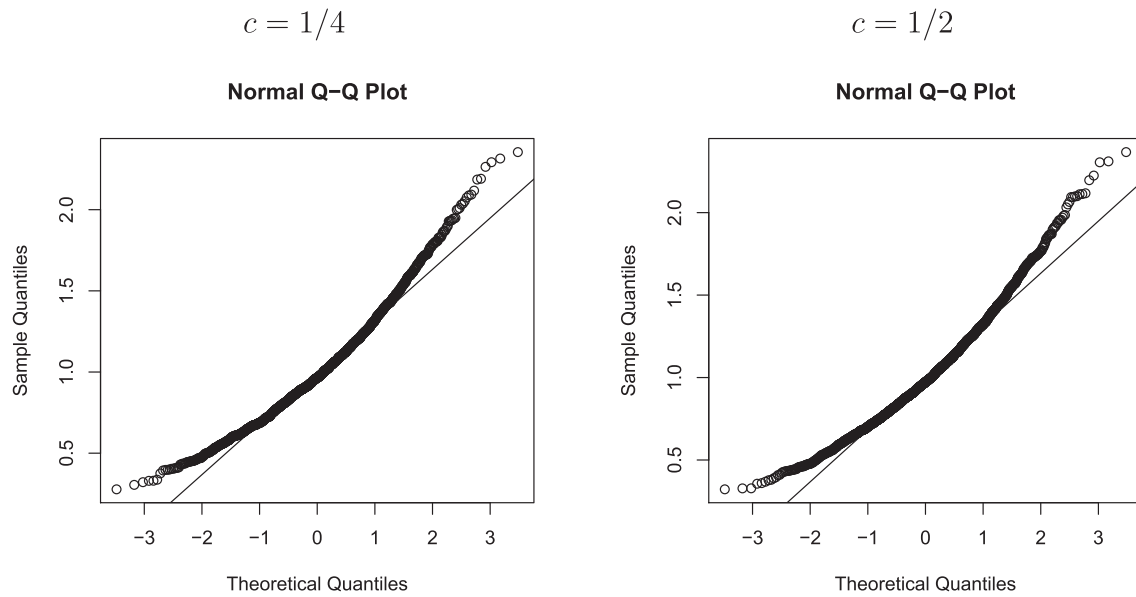
$$c = 1/4 \qquad\qquad c = 1/2$$

Figure 3. QQ plots for the estimates of $\gamma(1/2)$

Table 2. Rejection rates in simulations: sample size $n = 2000$.

			$k = 100$		$k = 50$	
		α	$h = 0.025$	$h = 0.040$	$h = 0.025$	$h = 0.040$
iid observations		0.1	0.089	0.079	0.132	0.079
		0.05	0.048	0.035	0.074	0.037
		0.01	0.011	0.004	0.022	0.007
Linear trend	$b = 1$	0.1	0.517	0.397	0.367	0.204
		0.05	0.406	0.285	0.260	0.124
		0.01	0.186	0.114	0.099	0.028
	$b = 2$	0.1	0.837	0.683	0.570	0.375
		0.05	0.760	0.542	0.461	0.266
		0.01	0.490	0.281	0.208	0.106
Sin trend	$c = 1/4$	0.1	0.277	0.438	0.222	0.227
		0.05	0.188	0.334	0.129	0.157
		0.01	0.071	0.138	0.038	0.040
	$c = 1/2$	0.1	0.876	0.945	0.582	0.696
		0.05	0.811	0.908	0.470	0.579
		0.01	0.574	0.765	0.228	0.304

4.2. Application 2: Loss Returns of S&P 500

We employ the same dataset as in Einmahl, de Haan, and Zhou (2016), that is, the S&P 500 index, from 1988 till 2012 . We construct daily loss returns defined as $X_t = \log(P_t/P_{t+1})$, where P_t is the index on day t. This results in a sample with 6302 observations.

Similar to Einmahl, de Haan, and Zhou (2016), we also test the constancy of the extreme value indices over a subperiod from 1988 to 2012 (6302 observations). The obtained p-values against various levels of k are shown in the *upper panel* of Figure 5. We do not reject the null hypothesis for k up to 750 under the 5% significance level (the dash line). This result differs from the conclusion in Einmahl, de Haan, and Zhou (2016) where the constancy of the extreme value index in the period from 1988 to 2012 was rejected. We interpret the difference by the fact that we are testing a broader null. Notice that in the heteroscedastic extremes model in Einmahl, de Haan, and Zhou (2016), the skedasis function is assumed to be bounded

away from 0 and $+\infty$, whereas our model potentially allows for unbounded skedasis in the null hypothesis. As a consequence, our test for the constancy of the extreme value index may have a lower power.

Since the null hypothesis is not rejected for the sample from 1988 to 2012, we consider an extended sample from 1963 to 2012 (12586 observations). The result is shown in the *lower panel* of Figure 5. We reject having a constant extreme value index during this long period for k ranging from 250 to 750 under the 5% confidence level. Based on this analysis, we conclude that there is a change of extreme value index during the period from 1963 to 2012.

One concern in the aforementioned analysis is that financial data such as stock returns exhibits serial dependence. The presence of serial dependence would in general enlarge the asymptotic variance of the local estimators for $\gamma(s)$. Correspondingly, the critical value of the proposed test should be higher. By using the test based on assuming no serial dependence, we tend to over
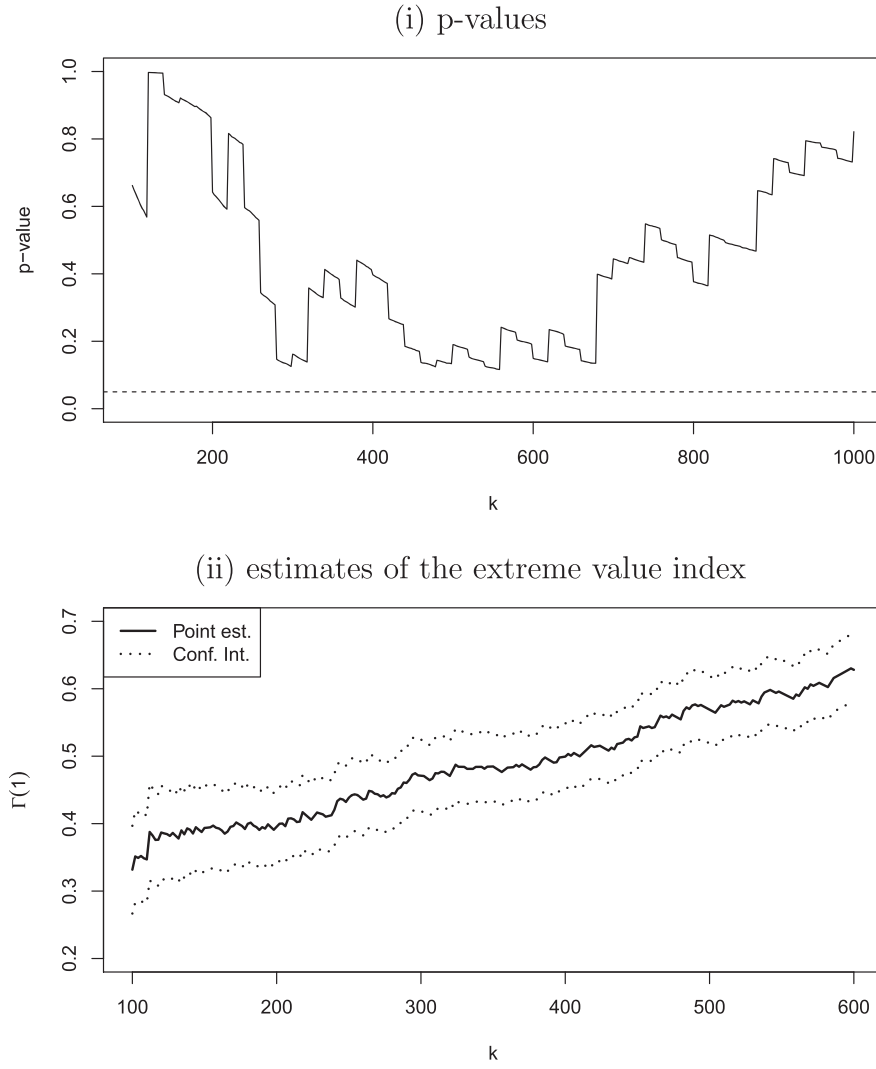
(i) p-values

(ii) estimates of the extreme value index

Figure 4. Testing the constancy of the extreme value indices: precipitation.

reject the null. Given that the analysis using the data from 1988 to 2012 did not reject the null, accounting for serial dependence may not alter the conclusion. However, the rejection result based on the data from 1963 to 2012 may suffer from the serial dependence issue. Therefore, we conduct additional analysis as follows. We split the dataset into two subsets that consist of daily returns on the even and odd days, respectively. In other words, we do not take returns from consecutive trading days. The split of the full dataset helps to mitigate the serial dependence and data from each subset is more close to the iid assumption. We conduct our tests in each subset with the results presented in Figure 6.

From the two figures in Figure 6, we observe that the null hypothesis is not rejected at 5% level, for the dataset containing daily returns on the even days only. For the daily returns on the odd days, the result is not conclusive either: for k ranging from 400 to 600, the null hypothesis is not rejected at 5% level. However for a lower choice of k, such as $k = 200$ or $k = 300$, the null is rejected under 5% level. Overall, the additional analysis reveals that the rejection result for the full dataset might be affected by the serial dependence. With accounting for serial dependence, there is no conclusive evidence that the extreme value index varies over this period.

Finally, we plot the estimated $\Gamma(s)$ function (the solid line) and the corresponding 95% confidence band uniformly for all $s \in [0, 1]$ in Figure 7 for the period from 1963 to 2012. In this analysis, we use $k = 400$, for which the null hypothesis of having a constant extreme value index was rejected. We obtain the confidence band in two ways.

Without having prior information on the shape of $\Gamma(s)$, we obtain from Theorem 2.2 that as $n \to \infty$,

$$\sup_{s \in [0,1]} \left| \sqrt{k} \left(\hat{\Gamma}_H(s) - \Gamma(s) \right) \right| \overset{d}{\to} \sup_{s \in [0,1]} \left| \int_0^s \gamma(u) dW(u) \right|,$$

where $W(u)$ is a Brownian motion. We simulate the quantile of the limit and use that for constructing the uniform confidence band. Since the limit distribution involves the function $\gamma(s)$, we plug the estimate of the $\gamma(s)$ function into the stochastic integral and simulate the statistic $\sup_{s \in [0,1]} \left| \int_0^s \hat{\gamma}_H(u) dW(u) \right|$ one million times. Then we take the numerical 95% quantile among the one million simulations, denoted as $q(0.95)$. Eventually, the uniform confidence band can be constructed as $[\hat{\Gamma}_H(s) - q(0.95)/\sqrt{k}, \hat{\Gamma}_H(s) + q(0.95)/\sqrt{k}]$. The upper and lower bound for the confidence band are shown by the dotted lines in the *upper panel* of Figure 7. The dash line indicates the function $s\hat{\Gamma}_H(1)$, which corresponds to the case $\gamma(s) = \gamma$ for all

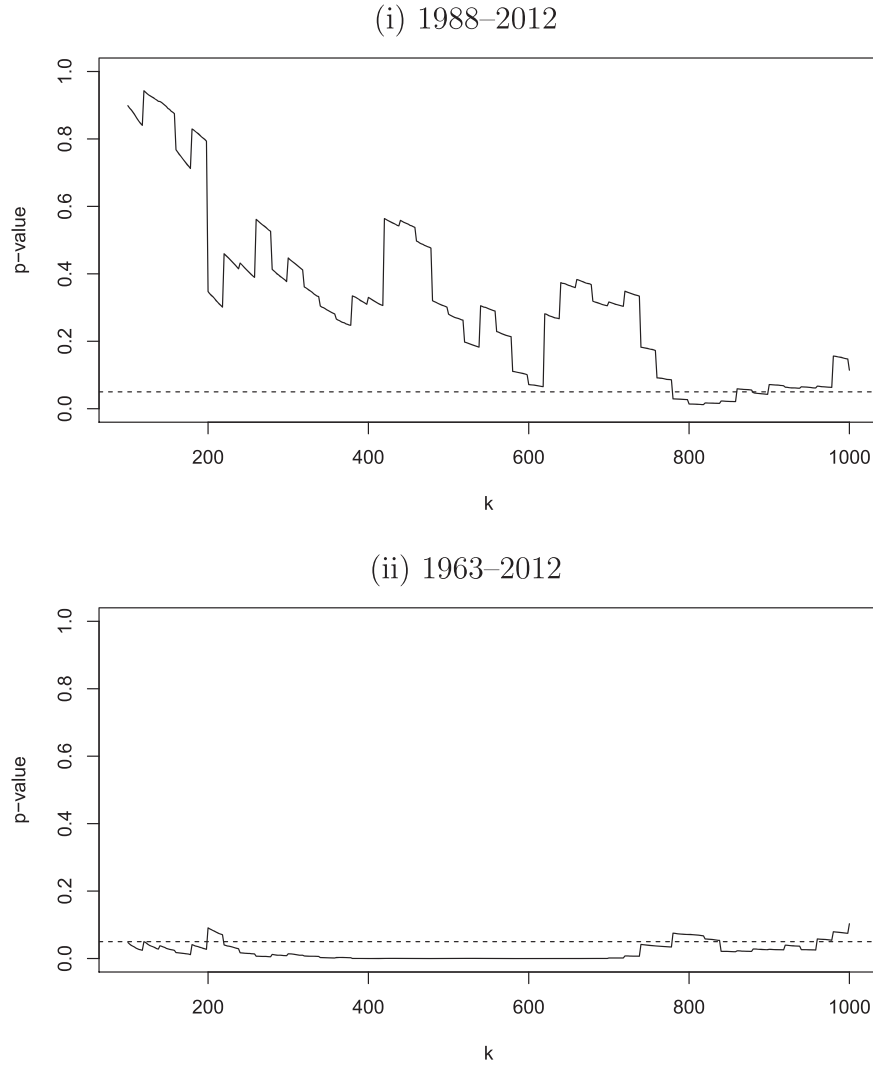## (i) 1988–2012

(ii) 1963–2012

Figure 5. Testing the constancy of the extreme value indices: S&P 500 index.

$s \in [0, 1]$. The dash line lies always within the confidence band which is seemingly contradictory to our testing result. Notice that the construction of the uniform confidence band is not based on the null hypothesis in the testing analysis. The width of the band is therefore relatively wider due to the stochastic integral.

To be consistent with the testing procedure, we construct a uniform confidence band under the null hypothesis that $\gamma(s) = \gamma$ for all $s \in [0, 1]$ in the *lower panel* of Figure 7. More specifically, we use the quantile of the limit distribution $\hat{\Gamma}_H(1) \sup_{s \in [0,1]} |B(s)|$ where $B(s)$ is a Brownian bridge, instead of the simulated $q(0.95)$ in determining the uniform width of the confidence band. The result shows that the straight line $s\hat{\Gamma}_H(1)$ falls out of the uniform confidence band in the region $s \in [0.3, 0.5]$. This is in line with our testing result.

Appendix A: Proofs

We start with presenting auxiliary results that are necessary for the proof of our main theorems. Then we establish the asymptotic theory for the "local tail empirical process." Finally, we provide the proof for the two main theorems.

A.1. Auxiliary Results

The following lemma shows that under the conditions of Theorem 2.1, the quantile function in the regularly variation property does not vary much in a h-neighborhood. Denote $q_n = k^{1+\varepsilon}$ for some $\varepsilon > 0$.

Lemma A.1. Under the conditions of Theorem 2.1, as $n \to \infty$,

$$\sqrt{k} \sup_{|s_1 - s_2| \leq h, 1/2 \leq x \leq q_n} \left| \frac{U_{s_1}\left(\frac{n}{k}x\right)}{U_{s_2}\left(\frac{n}{k}\right) x^{\gamma(s_2)}} - 1 \right| \to 0. \qquad (A.1)$$

Proof of Lemma A.1. Write

$$\frac{U_{s_1}\left(\frac{n}{k}x\right)}{U_{s_2}\left(\frac{n}{k}\right) x^{\gamma(s_2)}} = \frac{U_{s_1}\left(\frac{n}{k}x\right)}{U_{s_1}\left(\frac{n}{k}\right) x^{\gamma(s_1)}} \cdot \frac{U_{s_1}\left(\frac{n}{k}\right)}{U_{s_2}\left(\frac{n}{k}\right)} \cdot x^{\gamma(s_1) - \gamma(s_2)} =: I_1 \cdot I_2 \cdot I_3.$$

Firstly, the second order condition (3) ensures that (see Theorem 2.3.9 in de Haan and Ferreira (2006)) uniformly for all $s_1 \in [0, 1], x \geq 1/2$, as $n \to \infty$,

$$\left| \frac{I_1 - 1}{A_{s_1}\left(\frac{n}{k}\right)} \right| = O(1).$$
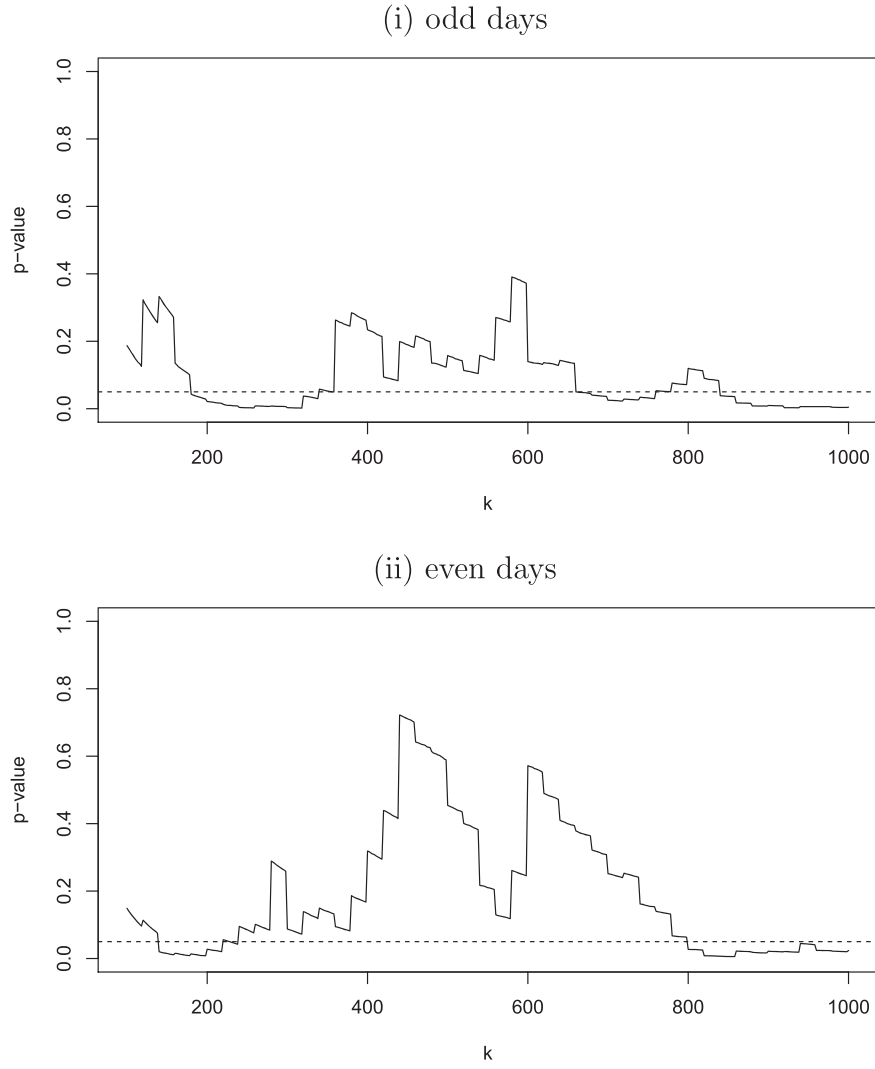
(i) odd days

(ii) even days

Figure 6. Testing the constancy of the extreme value indices: S&P 500 index (1963–2012).

Together with the condition (5), we get that as $n \to \infty$,

$$\sqrt{k} \sup_{s_1 \in [0,1], x \geq 1/2} |I_1 - 1| = o(1).$$

Secondly, the conditions (5) and (7) ensure that $n \to \infty$,

$$\sqrt{k} \sup_{|s_1 - s_2| \leq h} |I_2 - 1| = o(1).$$

Lastly, from the condition (6), we get that as $n \to \infty$,

$$\sqrt{k} \sup_{|s_1 - s_2| \leq h, 1/2 \leq x \leq q_n} |\log x| \, |\gamma(s_1) - \gamma(s_2)|$$

$$< 2\sqrt{k} \log k \sup_{|s_1 - s_2| \leq h} |\gamma(s_1) - \gamma(s_2)| \to 0.$$

It implies that $\sqrt{k} \sup_{|s_1 - s_2| \leq h, 1/2 \leq x \leq q_n} |I_3 - 1| = o(1)$.

The lemma is then proved by combining the three components. □

For any given $s \in [0,1]$, denote $\overline{U}_{s,n} = \max_{\{i: |\frac{i}{n} - s| \leq h\}} U_{\frac{i}{n}}$ and $\underline{U}_{s,n} = \min_{\{i: |\frac{i}{n} - s| \leq h\}} U_{\frac{i}{n}}$. Obviously $\underline{U}_{s,n} \leq U_s \leq \overline{U}_{s,n}$. The following is a direct corollary from Lemma A.1.

Corollary A.1. As $n \to \infty$,

$$\Delta_{4,n} := \sqrt{k} \sup_{0 \leq s \leq 1, 1/2 \leq x \leq q_n} \left| \frac{\overline{U}_{s,n}\left(\frac{n}{k}x\right)}{U_s\left(\frac{n}{k}\right)x^{\gamma(s)}} - 1 \right| \to 0,$$

$$\Delta_{5,n} := \sqrt{k} \sup_{0 \leq s \leq 1, 1/2 \leq x \leq q_n} \left| \frac{\underline{U}_{s,n}\left(\frac{n}{k}x\right)}{U_s\left(\frac{n}{k}\right)x^{\gamma(s)}} - 1 \right| \to 0.$$

A.2. Proof of Theorem 2.1

We prove the theorem by constructing upper and lower bounds for the local estimator $\hat{\gamma}_H(s)$ at a fixed s.

The local estimator is based on observations X_i with $i \in I_n(s) = \left\{i : \left|\frac{i}{n} - s\right| \leq h\right\}$. Write $X_i = U_{\frac{i}{n}}(Z_i)$ for $1 \leq i \leq n$, where $\{Z_i\}_{i=1}^n$ are iid standard Pareto distributed random variables. To construct the local Hill estimator, we rank the observations $\{X_i : i \in I_n(s)\}$ into order statistics as $X_{1,[2nh]}^{(s)} \leq \cdots \leq X_{[2nh],[2nh]}^{(s)}$. We also rank $\{Z_i : i \in I_n(s)\}$ into order statistics as $Z_{1,[2nh]}^{(s)} \leq \cdots \leq Z_{[2nh],[2nh]}^{(s)}$. Notice that since the $U_{\frac{i}{n}}$ functions are different, the order statistic $X_{j,[2nh]}^{(s)}$ may not correspond to the order statistic $Z_{j,[2nh]}^{(s)}$ for $j = 1, 2, \ldots, [2nh]$.
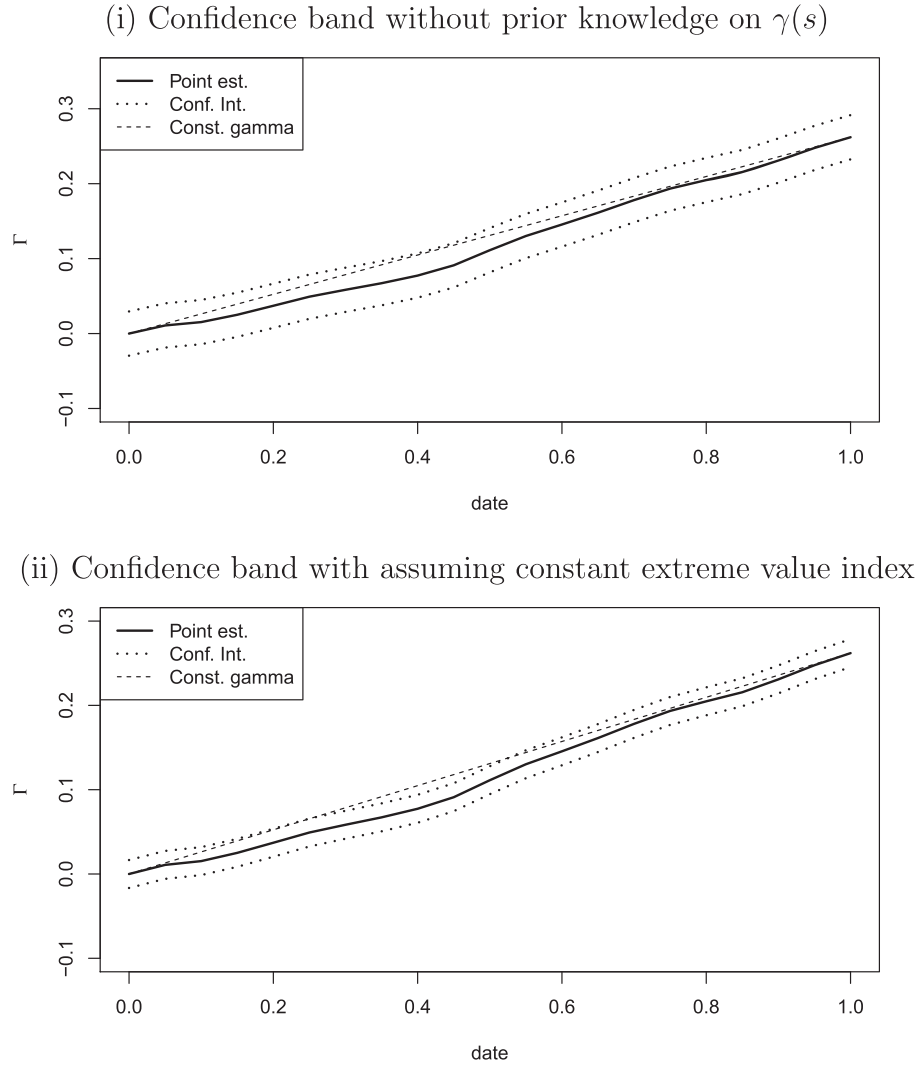
(i) Confidence band without prior knowledge on $\gamma(s)$

(ii) Confidence band with assuming constant extreme value index

Figure 7. Estimating the $\Gamma(s)$ function: S&P 500 index.

Nevertheless, recall the notations $\overline{U}_{s,n} = \max_{\{i:|\frac{i}{n}-s|\leq h\}} U_{\frac{i}{n}}$ and $\underline{U}_{s,n} = \min_{\{i:|\frac{i}{n}-s|\leq h\}} U_{\frac{i}{n}}$. The inequalities

$$\underline{U}_{s,n}(Z^{(s)}_{j,[2nh]}) \leq X^{(s)}_{j,[2nh]} \leq \overline{U}_{s,n}(Z^{(s)}_{j,[2nh]}),$$

for all $j = 1, 2, \ldots, [2nh]$, are obtained as follows. Take the upper bound as an example. Since there are j random variables among $\{Z_i : i \in I_n(s)\}$ that are lower or equal to $Z^{(s)}_{j,[2nh]}$, and the U-functions corresponding to these Z_i are all bounded below $\overline{U}$, we get that there are at least j random variables among $\{X_i : i \in I_n(s)\}$ that are bounded below by $\overline{U}_{s,n}(Z^{(s)}_{j,[2nh]})$. This proves the inequality for the upper bound. A similar argument can be made for the lower bound.

Therefore, we get an upper bound for $\hat{\gamma}_H(s)$ as follows,

$$\hat{\gamma}_H(s) = \frac{1}{[2kh]} \sum_{j=1}^{[2kh]} \left(\log X^{(s)}_{[2nh]-j+1,[2nh]} - \log X^{(s)}_{[2nh]-[2kh],[2nh]} \right)$$

$$\leq \frac{1}{[2kh]} \sum_{j=1}^{[2kh]} \left(\log \overline{U}_{s,n}(Z^{(s)}_{[2nh]-j+1,[2nh]}) \right.$$

$$\left. - \log \underline{U}_{s,n}(Z^{(s)}_{[2nh]-[2kh],[2nh]}) \right)$$

$$= \frac{1}{[2kh]} \sum_{j=1}^{[2kh]} \left(\log \frac{\overline{U}_{s,n}(Z^{(s)}_{[2nh]-j+1,[2nh]})}{U_s\left(\frac{n}{k}\right)} \right.$$

$$\left. - \log \frac{\underline{U}_{s,n}(Z^{(s)}_{[2nh]-[2kh],[2nh]})}{U_s\left(\frac{n}{k}\right)} \right). \tag{A.2}$$

We would now apply Corollary A.1 to bound the two terms in (A.2). For that purpose it is necessary to check that as $n \to \infty$,

$$\Pr\left(\frac{k}{n} Z^{(s)}_{[2nh]-j+1,[2nh]} \in [1/2, q_n], \text{ for all } j = 1, 2, \ldots, [2kh]\right) \to 1. \tag{A.3}$$

For the lower bound of $1/2$, it follows from the fact that $\frac{k}{n} Z^{(s)}_{[2nh]-[2kh],[2nh]} \xrightarrow{P} 1$, as $n \to \infty$. For the upper bound, notice that, as $n \to \infty$,

$$\Pr\left(\frac{k}{n} Z^{(s)}_{[2nh],[2nh]} \leq q_n, \text{ for all } s \in [0,1]\right)$$

$$\geq \Pr\left(\frac{k}{n} \max_{1\leq i\leq n} Z_i \leq k^{1+\varepsilon}\right) \geq \Pr\left(\frac{\max_{1\leq i\leq n} Z_i}{n} \leq k^\varepsilon\right) \to 1.$$

Here, we consider the maxima over all Z_i across all observations, and use the fact that $\frac{\max_{1\le i\le n} Z_i}{n} = O_p(1)$ as $n \to \infty$. In this way, we have verified the upper bound q_n in (A.3).

Now, we are ready to apply Corollary A.1 to the two terms in (A.2) and continue the inequality as follows: as $n \to \infty$, within the set $\left\{ \frac{k}{n} Z^{(s)}_{[2nh]-j+1,[2nh]} \in [1/2, q_n], \text{ for all } j = 1, 2, \ldots, [2kh] \right\}$,

$$
\hat{\gamma}_H(s) \le \frac{1}{[2kh]} \sum_{j=1}^{[2kh]} \left(\gamma(s) \log \frac{Z^{(s)}_{[2nh]-j+1,[2nh]}}{n/k} + \log\left(1 + \frac{\Delta_{4,n}}{\sqrt{k}}\right) \right.
$$
$$
\left. -\gamma(s) \log \frac{Z^{(s)}_{[2nh]-[2kh],[2nh]}}{n/k} - \log\left(1 - \frac{\Delta_{5,n}}{\sqrt{k}}\right) \right)
$$
$$
= \gamma(s) \frac{1}{[2kh]} \left(\sum_{j=1}^{[2kh]} \log \frac{Z^{(s)}_{[2nh]-j+1,[2nh]}}{Z^{(s)}_{[2nh]-[2kh],[2nh]}} \right)
$$
$$
+ \left(\log\left(1 + \frac{\Delta_{4,n}}{\sqrt{k}}\right) - \log\left(1 - \frac{\Delta_{5,n}}{\sqrt{k}}\right) \right)
$$
$$
=: \gamma(s) J_n(s) + \frac{1}{\sqrt{k}} o(1),
$$

where $J_n(s) = \frac{1}{[2kh]} \sum_{j=1}^{[2kh]} \log \frac{Z^{(s)}_{[2nh]-j+1,[2nh]}}{Z^{(s)}_{[2nh]-[2kh],[2nh]}}$. Note that the o(1) term in the remainder part $\frac{1}{\sqrt{k}} o(1)$ is uniform for all $s \in [0, 1]$.

Similarly, one can establish a lower bound for $\hat{\gamma}_H(s)$: as $n \to \infty$, within the set $\left\{ \frac{k}{n} Z^{(s)}_{[2nh]-j+1,[2nh]} \in [1/2, q_n], \text{ for all } j = 1, 2, \ldots, [2kh] \right\}$,

$$
\hat{\gamma}_H(s) \ge \gamma(s) J_n(s) + \frac{1}{\sqrt{k}} o(1).
$$

Hence, together with (A.3), we get that as $n \to \infty$,

$$
\hat{\gamma}_H(s) = \gamma(s) J_n(s) + \frac{1}{\sqrt{k}} o_P(1). \tag{A.4}
$$

To obtain the theorem, we only need to show that $\sqrt{2kh}(J_n(s) - 1)$ converges to a standard normal distribution as $n \to \infty$.

If we disregard the order, then the set $\left\{ \log \frac{Z^{(s)}_{[2nh]-j+1,[2nh]}}{Z^{(s)}_{[2nh]-[2kh],[2nh]}} : 1 \le j \le [2kh] \right\}$ is a set of iid standard exponentially distributed random variables. We denote them, without loss of generality, as $\left\{ E^{(s)}_j \right\}_{j=1}^{[2kh]}$. From the central limit theorem, we get that

$$
\sqrt{2kh}(J_n(s) - 1) = \sqrt{\frac{2kh}{[2kh]}} \sqrt{[2kh]} \left(\frac{1}{[2kh]} \sum_{j=1}^{[2kh]} \log E^{(s)}_j - 1 \right) \xrightarrow{d} N(0, 1),
$$

as $n \to \infty$. Notice that here we use the fact that $kh \to \infty$ as $n \to \infty$ from the condition (4).

A.3. Proof of Theorem 2.2

Next, we prove the global asymptotic normality in Theorem 2.2, which is a uniform result over all $s \in [0, 1]$.

Recall the definition of $\hat{\Gamma}_H(s)$ in (2) as

$$
\hat{\Gamma}_H(s) = \int_0^s \hat{\gamma}_H\left((2[\frac{u}{2h}] + 1)h \right) du,
$$

which is the partial integral of a discretized version of the function $\hat{\gamma}_H(\cdot)$. We apply the same discretization to the function $\gamma(\cdot)$ as follows: define

$$
\tilde{\gamma}(s) = \begin{cases} \gamma\left((2[\frac{s}{2h}] + 1)h \right), & \text{for } s \in [0, 2h[1/(2h)]), \\ \gamma\left((2[\frac{1}{2h}] - 1)h \right), & \text{for } s \in [2h[1/(2h)], 1]. \end{cases}
$$

Further define $\tilde{\Gamma}(s) = \int_0^s \tilde{\gamma}(u) du$. We first compare $\hat{\Gamma}_H(s)$ to $\tilde{\Gamma}(s)$ and show that under a Skorokhod construction there exists a series of Brownian motions $\{W_n(s)\}$ such that as $n \to \infty$,

$$
\sup_{s\in[0,1]} \left| \sqrt{k}\left(\hat{\Gamma}_H(s) - \tilde{\Gamma}(s) \right) - \int_0^s \tilde{\gamma}(u) dW_n(u) \right| \xrightarrow{P} 0. \tag{A.5}
$$

Later on, we handle the uniformly negligible difference between $\tilde{\gamma}(\cdot)$ and $\gamma(\cdot)$.

Since the estimator $\hat{\Gamma}_H(s)$ involves the local estimators $\hat{\gamma}_H(s)$ at various locations $s = (2p-1)h$ for $p = 1, 2, \ldots, [1/2h]$, we first revisit the proof of Theorem 2.1 to get the expansion (A.4) uniformly for all $s = (2p-1)h, p = 1, 2, \ldots, [1/(2h)]$. One important step in the proof of Theorem 2.1 is to verify that $\Pr\left(\frac{1}{2} \le \frac{k}{n} Z^{(s)}_{[2nh]-j+1,[2nh]} \le q_n \right) \to 1$ as $n \to \infty$, for a fixed s. To achieve the uniform result, we have to show that this inequality holds uniformly for all $s = (2p-1)h$, $p = 1, 2, \ldots, [1/(2h)]$. The following lemma guarantees the validity of such a uniform inequality.

Lemma A.2. Under the conditions in Theorem 2.2, as $n \to \infty$,

$$
\Pr\left(\frac{k}{n} Z^{((2p-1)h)}_{[2nh]-j+1,[2nh]} \in [1/2, q_n], \forall 1 \le j \le [2kh] + 1, \right.
$$
$$
\left. 1 \le p \le [1/(2h)] \right) \to 1.
$$

Proof of Lemma A.2. The validity of the upper bound q_n has already been proved in the proof of Theorem 2.1 as follows. As $n \to \infty$, $\Pr\left(\frac{k}{n} Z^{(s)}_{[2nh],[2nh]} \le q_n, \text{ for all } s \right) \to 1$. Consequently, it holds for all $s = (2p-1)h, p = 1, 2, \ldots, [1/(2h)]$.

For the lower bound, we shall prove that: as $n \to \infty$, $\Pr\left(\frac{k}{n} Z^{((2p-1)h)}_{[2nh]-[2kh],[2nh]} \ge 1/2, \text{ for all } p = 1, 2, \ldots, [1/(2h)] \right) \to 1$. Notice that the order statistics $Z^{((2p-1)h)}_{[2nh]-[2kh],[2nh]}$ are iid across $p = 1, 2, \ldots, [1/(2h)]$. This implies that, for any fixed s_0,

$$
\Pr\left(\frac{k}{n} Z^{((2p-1)h)}_{[2nh]-[2kh],[2nh]} \ge 1/2, \forall 1 \le p \le [1/(2h)] \right)
$$
$$
= \left(\Pr\left(\frac{k}{n} Z^{(s_0)}_{[2nh]-[2kh],[2nh]} \ge 1/2 \right) \right)^{[1/(2h)]}.
$$

We continue to establish the lower bound for the right hand side. For convenience, we omit the superscript (s_0). Since $Z_{[2nh]-[2kh],[2nh]}$ is the $([2nh] - [2kh])$th order statistic among $[2nh]$ iid standard Pareto distributed random variables, $1/Z_{[2nh]-[2kh],[2nh]}$ is the $([2kh] + 1)$th order statistic, $U_{[2kh]+1,[2nh]}$, among $[2nh]$ iid uniformly distributed random variables $U_1, U_2, \ldots, U_{[2nh]}$. Hence,

$$
\Pr\left(\frac{k}{n} Z_{[2nh]-[2kh],[2nh]} \ge 1/2 \right) = 1 - \Pr\left(\frac{n}{k} U_{[2kh]+1,[2nh]} > 2 \right)
$$
$$
= 1 - \Pr\left(\sum_{i=1}^{[2nh]} 1_{\{U_i \le 2k/n\}} \le [2kh] \right)
$$
$$
\ge 1 - \exp\left\{ -\frac{[2nh]}{2} \left(\frac{[2kh]}{[2nh]} - \frac{2k}{n} \right)^2 \frac{1}{\frac{2k}{n}\left(1 - \frac{2k}{n}\right)} \right\}
$$
$$
\ge 1 - \exp(-Ckh),
$$

for some $C > 0$. Here the first inequality is an application of a stronger version of the Hoeffding's inequality for binomial distributions: Theorem 2(ii) in Okamoto (1959), with replacing n, p, and c therein by $[2nh]$, $\frac{2k}{n}$ and $\frac{2k}{n} - \frac{[2kh]}{[2nh]}$, respectively. Consequently,

$$\Pr\left(\frac{k}{n}Z^{((2p-1)h)}_{[2nh]-[2kh],[2nh]} \geq 1/2 \text{ for all } p = 1, 2, \ldots, [1/(2h)]\right)$$
$$\geq \left(1 - \exp(-Ckh)\right)^{[1/(2h)]}.$$

As $n \to \infty$, we have that $\frac{1}{2h}\exp(-Ckh) \to 0$ since $kh/\left|\log h\right| \to \infty$ and $kh \to \infty$, see the condition (4). The lemma is proved. □

We continue with the proof of Theorem 2.2. Recall that in the proof of Theorem 2.1, the expansion (A.4) is proved for each given s. From the uniform bounds for all $s = (2p-1)h$, $p = 1, 2, \ldots, [1/(2h)]$ given in Lemma A.2, following the same lines, we obtain the expansion (A.4) uniformly for all $s = (2p-1)h$, $p = 1, 2, \ldots, [1/(2h)]$, that is, as $n \to \infty$,

$$\sqrt{k} \sup_{1 \leq p \leq [1/(2h)]} \left|\hat{\gamma}_H((2p-1)h) - \gamma((2p-1)h)J^{(p)}_n\right| = o_P(1),$$

where $J^{(p)}_n = \frac{1}{[2kh]}\sum_{j=1}^{[2kh]}\log\frac{Z^{((2p-1)h)}_{[2nh]-j+1,[2nh]}}{Z^{((2p-1)h)}_{[2nh]-[2kh],[2nh]}}$.

Again, we regard the set $\left\{\log\frac{Z^{((2p-1)h)}_{[2nh]-j+1,[2nh]}}{Z^{((2p-1)h)}_{[2nh]-[2kh],[2nh]}} : 1 \leq j \leq [2kh]\right\}$ as a set of iid standard exponentially distributed random variables denoted as $\left\{E^{(p)}_j : 1 \leq j \leq [2kh]\right\}$. Then $J^{(p)}_n = \frac{1}{[2kh]}\sum_{j=1}^{[2kh]}E^{(p)}_j$.

We deal with the two regions $s \leq 2h[1/(2h)]$ and $s \geq 2h[1/(2h)]$ separately.

For $s \leq 2h[1/(2h)]$, we obtain the expansion of the estimator $\hat{\Gamma}_H(s)$ as follows:

$$\hat{\Gamma}_H(s) = 2h\sum_{p=1}^{[s/(2h)]}\hat{\gamma}_H((2p-1)h) + (s - 2h[s/(2h)])\hat{\gamma}_H\left((2[s/(2h)]+1)h\right).$$

Notice that

$$\tilde{\Gamma}_H(s) = 2h\sum_{p=1}^{[s/(2h)]}\gamma((2p-1)h) + (s - 2h[s/(2h)])\gamma\left((2[s/(2h)]+1)h\right).$$

By comparing the two with applying the uniform asymptotic expansion of $\hat{\gamma}_H((2p-1)h)$, for $p = 1, 2, \ldots, [s/(2h)] + 1$, we get that

$$\sup_{s \leq 2h[1/(2h)]}\left|\sqrt{k}\left(\hat{\Gamma}_H(s) - \tilde{\Gamma}(s)\right) - \left(L^{(1)}_n(s) + L^{(2)}_n(s)\right)\right| = o_P(1),$$

where

$$L^{(1)}_n(s) := 2h\sqrt{k}\sum_{p=1}^{[s/(2h)]}\gamma_H((2p-1)h)(J^{(p)}_n - 1),$$

$$L^{(2)}_n(s) := \sqrt{k}(s - 2h[s/(2h)])$$
$$\gamma_H\left((2[s/(2h)]+1)h\right)(J^{([s/(2h)]+1)}_n - 1).$$

We further study the asymptotic behavior of $L^{(1)}_n(s)$ and $L^{(2)}_n(s)$.

Notice that for any $1 \leq p \neq q \leq [1/(2h)]$, the interval $[(2p-2)h, 2ph)$ and $[(2q-2)h, 2qh)$ are disjoint. Therefore, we can collect all $E^{(p)}_j$ for all $j = 1, 2, \ldots, [2kh]$ and $p = 1, 2, \ldots, [1/(2h)]$ and regard them as $[2kh][1/(2h)]$ iid standard exponentially distributed random variables. The random variables $E^{(p)}_j$ for $j = 1, 2, \ldots, [2kh]$ and

$p = 1, 2, \ldots, [1/(2h)]$ is a sequence of $m = [2kh][1/(2h)]$ independent standard exponential random variables. For convenience, we relabel $E^{(p)}_j - 1$ as $Y_1, Y_2, \ldots, Y_m$, that is, $Y_{(p-1)[2kh]+j} = E^{(p)}_j - 1$ and define the partial sum process

$$S_n(u) = \frac{1}{\sqrt{m}}\sum_{i=1}^{[mu]}Y_i,$$

for $u \in [0, 1]$. With a variation of the notation

$$\tilde{S}_n(u) := S_n\left(\frac{u}{2h[1/(2h)]}\right),$$

we can rewrite that

$$J^{(p)}_n - 1 = \frac{1}{[2kh]}\sum_{j=1}^{[2kh]}(E^{(p)}_j - 1) = \frac{\sqrt{m}}{[2kh]}(\tilde{S}_n(2hp) - \tilde{S}_n(2h(p-1))),$$

for $p = 1, 2, \ldots, [1/(2h)]$. With this new notation, we have that

$$L^{(1)}_n(s) = \frac{2h\sqrt{km}}{[2kh]}\int_0^{2[s/(2h)]h}\tilde{\gamma}(u)d\tilde{S}_n(u),$$

$$L^{(2)}_n(s) = \frac{2h\sqrt{km}}{[2kh]}\gamma((2[s/(2h)]+1)h)\left(\frac{s}{2h} - \left[\frac{s}{2h}\right]\right)$$
$$\int_{2h[s/(2h)]}^{2h[s/(2h)]+2h}d\tilde{S}_n(u).$$

To handle the $\tilde{S}_n$ process, we first apply Theorem 2.2(ii) in Csörgő and Horváth (1993) to get an asymptotic expansion of the S_n process. Note that $E(e^{tY_1}) = e^{-t}\frac{1}{1-t}$ for all $t < 1$. By verifying the condition in that theorem, we obtain that under a Skorokhod construction, there exists a series of Brownian motions $W_n(u)$ and a constant C_1 such that

$$S_n(u) = W_n(u) + \theta_n(u),$$

where as $n \to \infty$,

$$\Pr\left(\sup_{u \in [0,1]}|\theta_n(u)| \leq C_1\frac{\log m}{\sqrt{m}}\right) \to 1. \tag{A.6}$$

Define

$$\tilde{\theta}_n(u) := \tilde{S}_n(u) - W_n(u) = W_n\left(\frac{u}{2h[1/(2h)]}\right)$$
$$- W_n(u) + \theta_n\left(\frac{u}{2h[1/(2h)]}\right).$$

As $n \to \infty$, from the modulus of continuity, we have that

$$\Pr\left(\sup_{u \in [0,1]}\left|W_n\left(\frac{u}{2h[1/(2h)]}\right) - W_n(u)\right|\right.$$
$$\left.\leq C_2(2h)^{1/2}\left|\log(2h)\right|^{1/2}\right) \to 1,$$

for some positive constant C_2. Together with (A.6), we get that, as $n \to \infty$,

$$\Pr\left(\sup_{u \in [0,2h[1/(2h)]]}\left|\tilde{\theta}_n(u)\right|\right.$$
$$\left.\leq C_2(2h)^{1/2}\left|\log(2h)\right|^{1/2} + C_1\frac{\log m}{\sqrt{m}}\right) \to 1,$$

which implies that

$$\Pr\left(\sup_{u\in[0,2h[1/(2h)]]}\left|\tilde{\theta}_n(u)\right|\right.$$
$$\left.\leq C_3(h^{1/2}\left|\log h\right|^{1/2} + k^{-1/2}\log k)\right) \to 1, \qquad (A.7)$$

for some positive constants C_3.

Write, for $j = 1, 2$,

$$L_n^{(j)}(s) = L_n^{(W,j)}(s) + L_n^{(\tilde{\theta},j)}(s),$$

where $L_n^{(W,j)}(s)$ and $L_n^{(\tilde{\theta},j)}(s)$ are defined by replacing the process $\tilde{S}_n$ in the expression of $L_n^{(j)}$ by the processes W_n and $\tilde{\theta}_n$, respectively. Since $\tilde{\gamma}$ is a step function, $L_n^{(W,j)}(s)$ and $L_n^{(\tilde{\theta},j)}(s)$ should be read as sums instead of stochastic integrals.

Firstly, we handle $L_n^{(W,1)}(s)$ and $L_n^{(W,2)}(s)$ as follows. Notice that

$$\frac{[2kh]}{2h\sqrt{km}}L_n^{(W,1)}(s) = \int_0^{2[s/(2h)]h}\tilde{\gamma}(u)dW_n(u)$$
$$= \int_0^s \tilde{\gamma}(u)dW_n(u)$$
$$- \gamma((2[s/(2h)]+1)h)\int_{2h[s/(2h)]}^s dW_n(u).$$

With denoting $w(s,h) = \frac{s}{2h} - \left[\frac{s}{2h}\right] \in [0,1)$, we have that as $n \to \infty$,

$$\left|\frac{[2kh]}{2h\sqrt{km}}\left(L_n^{(W,1)}(s) + L_n^{(W,2)}(s)\right) - \int_0^s \tilde{\gamma}(u)dW_n(u)\right|$$
$$= \gamma((2[s/(2h)]+1)h)\left|w(s,h)\int_{2h[s/(2h)]}^{2h[s/(2h)]+2h} dW_n(u)\right.$$
$$\left.- \int_{2h[s/(2h)]}^s dW_n(u)\right|$$
$$= \gamma((2[s/(2h)]+1)h)\left|w(s,h)\int_s^{2h[s/(2h)]+2h} dW_n(u)\right.$$
$$\left.- (1 - w(s,h))\int_{2h[s/(2h)]}^s dW_n(u)\right|$$
$$\leq \tilde{\gamma}\left(\left|\int_s^{2h[s/(2h)]+2h} dW_n(u)\right| + \left|\int_{2h[s/(2h)]}^s dW_n(u)\right|\right)$$
$$= o_P(1),$$

where $\tilde{\gamma}$ is the uniform upper bound of the $\gamma(\cdot)$ function and the final $o_P(1)$ term is uniformly for all $s \leq 2h[1/(2h)]$ thanks to the uniform continuity of the Brownian motion. Since $\frac{[2kh]}{2h\sqrt{km}} \to 1$ as $n \to \infty$, we get that

$$\sup_{s\leq 2h[1/(2h)]}\left|L_n^{(W,1)}(s) + L_n^{(W,2)}(s) - \int_0^s \tilde{\gamma}(u)dW_n(u)\right| = o_P(1).$$

Next, we handle $L_n^{(\tilde{\theta},j)}(s)$ for $j = 1, 2$ by inequalities. As $n \to \infty$, $\frac{[2kh]}{2h\sqrt{km}} \to 1$, which leads to

$$\left|L_n^{(\tilde{\theta},1)}(s) + L_n^{(\tilde{\theta},2)}(s)\right|$$
$$\leq 2\left|\sum_{p=1}^{[s/(2h)]}\gamma((2p-1)h)(\tilde{\theta}_n(2ph) - \tilde{\theta}_n(2(p-1)h))\right.$$
$$+ \gamma((2[s/(2h)]+1)h)w(s,h) \cdot$$
$$\left.(\tilde{\theta}_n(2h[s/(2h)]+2h) - \tilde{\theta}_n(2h[s/(2h)]))\right|$$
$$\leq 2\left|\sum_{p=1}^{[s/(2h)]}\tilde{\theta}_n(2ph)(\gamma((2p-1)h) - \gamma((2p+1)h))\right|$$
$$+ 2\tilde{\gamma}\left(\left|\tilde{\theta}_n(2h[s/(2h)]+2h)\right| + 2\left|\tilde{\theta}_n(2h[s/(2h)])\right|\right)$$
$$\leq 2[s/2h]\sup_{u\in[0,1]}\left|\tilde{\theta}_n(u)\right| \cdot \frac{\Delta_{2,n}}{\sqrt{k}\log k} + 6\tilde{\gamma}\sup_{u\in[0,1]}\left|\tilde{\theta}_n(u)\right|$$
$$\leq \left(\frac{2}{2h\sqrt{k}\log k} + 6\tilde{\gamma}\right)C_3\left(h^{1/2}\left|\log h\right|^{1/2} + k^{-1/2}\log k\right),$$

where $\tilde{\gamma}$ is the uniform upper bound of the $\gamma(\cdot)$ function. In the last step, we use the inequality (A.7). The conditions (4) and (6) imply that as $n \to \infty$, the upper bound, which does not depend on s, converges to zero. Therefore, by combining the four components, $L_n^{(W,j)}(s)$ and $L_n^{(\tilde{\theta},j)}(s)$ for $j = 1, 2$, we get that as $n \to \infty$,

$$\sup_{s\leq 2h[1/(2h)]}\left|\sqrt{k}\left(\hat{\Gamma}_H(s) - \tilde{\Gamma}(s)\right) - \int_0^s \tilde{\gamma}(s)dW_n(s)\right| = o_P(1).$$

We extend this result to include the remaining part $s \in [2h[1/(2h)], 1]$. We first apply the established result for $s = 2h[1/(2h)]$ to obtain that as $n \to \infty$,

$$\left|\sqrt{k}\left(\hat{\Gamma}_H(2h[1/(2h)]) - \tilde{\Gamma}(2h[1/(2h)])\right) - \int_0^{2h[1/(2h)]}\tilde{\gamma}(s)dW_n(s)\right|$$
$$= o_P(1).$$

Therefore, to include the remaining part $s \in [2h[1/(2h)], 1]$, it suffices to show that as $n \to \infty$,

$$\sup_{s\in[2h[1/(2h)],1]}\sqrt{k}(s - 2h[1/(2h)])\left|\hat{\gamma}_H((2[1/(2h)]-1)h)\right.$$
$$\left.- \gamma((2[1/(2h)]-1)h)\right| = o_P(1),$$
$$\sup_{s\in[2h[1/(2h)],1]}\left|\gamma((2[1/(2h)]-1)h)(W_n(s) - W_n(2h[1/(2h)]))\right|$$
$$= o_P(1).$$

The first limit relation is ensured by the local asymptotic normality of $\hat{\gamma}_H((2[1/(2h)]-1)h)$. The second limit relation follows from the uniform continuity of W_n and the fact that the $\gamma(\cdot)$ function is uniformly bounded. Therefore, we can extend the region of s to the full interval $[0, 1]$ and obtain (A.5).

Finally, we show that $\tilde{\gamma}(\cdot)$ and $\tilde{\Gamma}(\cdot)$ can be replaced by $\gamma(\cdot)$ and $\Gamma(\cdot)$ in (A.5). From the definition of $\tilde{\gamma}(\cdot)$, it is straightforward to verify that

$$\sup_{s\in[0,1]}|\tilde{\gamma}(s) - \gamma(s)| \leq \sup_{|s_1-s_2|\leq 2h}|\gamma(s_1) - \gamma(s_2)| = \frac{\Delta_{2,n}}{\sqrt{k}\log k}.$$

Consequently, the condition (6) ensures that as $n \to \infty$,

$$\sup_{s\in[0,1]}\sqrt{k}\left|\int_0^s \gamma(u)du - \int_0^s \tilde{\gamma}(u)du\right| \leq \frac{\Delta_{2,n}}{\log k} \to 0.$$

Lastly, we show that as $n \to \infty$,

$$\sup_{s \in [0,1]} \left| \int_0^s \gamma(u) dW_n(u) - \int_0^s \tilde{\gamma}(u) dW_n(u) \right| \xrightarrow{P} 0. \quad (A.8)$$

Notice that for any fixed n, $\left\{ \int_0^s (\tilde{\gamma}(u) - \gamma(u)) dW_n(u) \right\}_{0 \le s \le 1}$ is a martingale. We apply the Doob's inequality for the sub-martingale $\left\{ \left| \int_0^s (\tilde{\gamma}(u) - \gamma(u)) dW_n(u) \right| \right\}_{0 \le s \le 1}$ and get that for any fixed n and $\varepsilon > 0$,

$$\Pr\left(\sup_{s \in [0,1]} \left| \int_0^s \gamma(u) dW_n(u) - \int_0^s \tilde{\gamma}(u) dW_n(u) \right| \ge \varepsilon \right)$$

$$\le \frac{1}{\varepsilon^2} E \left| \int_0^1 \gamma(u) dW_n(u) - \int_0^s \tilde{\gamma}(u) dW_n(u) \right|^2$$

$$= \frac{1}{\varepsilon^2} \int_0^1 (\gamma(u) - \tilde{\gamma}(u))^2 du \le \frac{1}{\varepsilon^2} \left(\frac{\Delta_{2,n}}{\sqrt{k} \log k} \right)^2.$$

By taking $n \to \infty$ and applying the condition (6), we obtain (A.8). Consequently, we have proved the theorem.

Acknowledgments

References

Balkema, A., and de Haan, L. (1974), "Residual Life Time at Great Age," *Annals of Probability*, 2, 792–804. [1]

Csörgő, M., and Horváth, L. (1993), *Weighted Approximations in Probability and Statistics*, New York: Wiley. [13]

Davison, A., and Ramesh, N. (2000), "Local Likelihood Smoothing of Sample Extremes," *Journal of the Royal Statistical Society*, Series B, 62, 191–208. [2]

Davison, A., and Smith, R. (1990), "Models for Exceedances Over High Thresholds," *Journal of the Royal Statistical Society*, Series B, 52, 393–442. [2]

de Haan, L., and Ferreira, A. (2006), *Extreme Value Theory: An Introduction*, New York: Springer. [9]

Einmahl, J., de Haan, L., and Zhou, C. (2016), "Statistics of Heteroscedastic Extremes," *Journal of the Royal Statistical Society*, Series B, 78, 31–51. [2,4,7]

Einmahl, J., and Lin, T. (2006), "Asymptotic Normality of Extreme Value Estimators on C[0, 1]," *Annals of Statistics*, 34, 469–492. [3]

Gardes, L., and Girard, S. (2010), "Conditional Extremes From Heavy-Tailed Distributions: An Application to the Estimation of Extreme Rainfall Return Levels," *Extremes*, 13, 177–204. [2]

Goegebeur, Y., Guillou, A., and Schorgen, A. (2014), "Nonparametric Regression Estimation of Conditional Tails: The Random Covariate Case," *Statistics*, 48, 732–755. [2]

Hall, P., and Tajvidi, N. (2000), "Nonparametric Analysis of Temporal Trend When Fitting Parametric Models to Extreme Value Data," *Statistical Science*, 15, 153–167. [2]

Hoga, Y. (2017), "Testing for Changes in (Extreme) VaR," *Econometrics Journal*, 20, 23–51. [2]

Okamoto, M. (1959), "Some Inequalities Relating to the Partial Sum of Binomial Probabilities," *Annals of the Institute of Statistical Mathematics*, 10, 29–35. [13]

Quintos, C., Fan, Z., and Phillips, P. C. (2001), "Structural Change Tests in Tail Behaviour and the Asian Crisis," *Review of Economic Studies*, 68, 633–663. [2]

Smith, R. L. (1989), "Extreme Value Analysis of Environmental Time Series: An Application to Trend Detection in Ground-Level Ozone," *Statistical Science*, 4, 367–377. [2]