



HAL
open science

Improved estimators of tail index and extreme quantiles under dependence serials

Mamadou Aliou, Hadji El, Aba Diop, Solym Mawaki Manou-Abi

► **To cite this version:**

Mamadou Aliou, Hadji El, Aba Diop, Solym Mawaki Manou-Abi. Improved estimators of tail index and extreme quantiles under dependence serials. 2023. hal-03968501

HAL Id: hal-03968501

<https://inria.hal.science/hal-03968501v1>

Preprint submitted on 1 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Improved estimators of tail index and extreme quantiles under dependence serials

Mamadou Aliou Barry ⁽¹⁾ , El Hadji Deme, ^(1,*)
Aba Diop ⁽²⁾ and Solym M. Manou-Abi ⁽³⁾

(1) Laboratoire d'Etudes et de Recherche en Statistiques et Développement, Saint-Louis, Sénégal

(2) Département de Mathématiques, Université Alioune Diop de Bambey, Sénégal (Equipe de recherche en Statistique et Modeles Aléatoires (ERESMA

(3) Institut Montpelliérain Alexander Grothendieck, UMR CNRS 5149, Place Eugène Bataillon, 34090, Montpellier, France

Abstract In this paper, we deal with the estimation problem for the extreme value parameters in the case of stationary β -mixing serials with heavy-tailed distributions. We first introduce two families of estimators generalizing the Hill's estimator. And from those families, three (03) asymptotically unbiased estimators of the extreme value index are established. Our reflection is based on the generalized Jackknife methodology which consists of taking any pair of three special cases of our family of estimators to cancel the bias term. The resulting estimators are also used to deduce three asymptotically unbiased estimators of the extreme quantiles. In a simulation survey, the performance of our proposed methods are compared to alternative estimators recently introduced in the literature. Finally, our methods are applied to high financial losses data in order to estimate the Value-at-Risk of the daily stock returns on the S&P500 index.

Keywords: Estimation, Tail index, Extreme quantiles, Heavy-tailed, Bias reduction, Dependent serials

(*)Corresponding author: elhadji.deme@ugb.edu.sn

1. Introduction

Severe damages and losses often occur when unusual and intense events happen. For instance, extremely high water levels may cause a dike to smash and also large negative stocks return can produce a high financial losses or bankruptcy. Therefore, it is important to model these potential losses or the corresponding return periods in order to analyze and evaluate them. The Extreme Value Theory (EVT) helps to depict such events by estimating the parameter of rare events, i.e, tail risks.

The independence and identically distributed assumption (i.i.d.) is primordial for classical EVT models. With such models, our main interest is on the estimation of the tail index parameter which is the basic parameter in extreme value theory. Several approaches for estimating this parameter have been proposed, including the famous Hill's estimator ((17)). And its generalization has been investigated in that regard ((12)). However, some data such as financial or environmental data expose serial dependence. This is the motivation of (2), (5), and (16) to adapt some estimators in the dependence case.

More precisely, in their work, they assumed that (X_1, X_2, \dots) is a β -mixing time series, that is, a series such that

$$\beta(m) := \sup_{p \geq 1} \mathbb{E} \left\{ \sup_{C \in \mathcal{B}_{p+m}^\infty} |\mathbb{P}(C | \mathcal{B}_1^p) - \mathbb{P}(C)| \right\} \rightarrow 0,$$

as $m \rightarrow \infty$, where \mathcal{B}_i^j denotes the σ -algebra generated by (X_i, \dots, X_j) . Without loss of generality, $\beta(m)$ measures the total variation distance between the unconditional distribution of the future of the time series and the conditional distribution of the future given the past of the time series when both are detached by m time points. Let F be the common marginal distribution function of $X_i, i \in \mathbb{N}$, assuming to be heavy-tailed (belonging to the the Fréchet domain of attraction), that is, there exist a positive number γ and the tail quantile function $U := (1/1 - F)^\leftarrow$ where $^\leftarrow$ denotes the left-continuous inverse function, such that

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\gamma, \forall x > 0. \quad (1.1)$$

From the relation in (1.1), U is said to be a regularly varying function at infinity with index γ . The parameter γ is called tail index or extreme value index and controls the behavior of the tail distribution function. Its estimation has mostly been studied in the case of i.i.d. random variables, although only few papers consider that for the case of time series with serial dependence features. We can mention among others, (5), (6) and (18). And very recently (2) and (16) are explored. Furthermore, in the context of i.i.d. assumption, the simplest estimator for $\gamma > 0$ is the Hill's estimator (17) defined by

$$\hat{\gamma}_{k_n}^{(H)} := \frac{1}{k_n} \sum_{i=1}^{k_n} \log X_{n-i+1,n} - \log X_{n-k_n,n},$$

where $X_{1,n}, \leq \dots \leq X_{n,n}$ stands for the order statistics and k_n represents an intermediate sequence, that is, a sequence such that $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$, as $n \rightarrow \infty$. To prove the asymptotic normality of the tail index estimators such as the Hill's one, we need a second order condition which specifies the rate of convergence for the left-hand side in (1.1) to its limit. This condition can be formulated in different ways as shown below. We will use the formulation later-on.

Second order condition (C_{SO}). Suppose that there exists a positive or negative function A with $\lim_{t \rightarrow \infty} A(t) = 0$ and a real number $\rho < 0$ such that

$$\lim_{t \rightarrow \infty} \frac{1}{A(t)} \left(\frac{U(tx)}{U(t)} - x^\gamma \right) = x^\gamma \frac{x^\rho - 1}{\rho}, \quad \forall x > 0. \quad (1.2)$$

The rate of the convergence for the function A to 0 is essential since it helps to exhibit the bias term of the tail index estimators. From the assumption that the intermediate sequence k_n is such that $k_n^{1/2} A(n/k_n) \rightarrow \lambda \in \mathbb{R}$, as $n \rightarrow \infty$ and assuming the following regularity conditions on the β -mixing coefficients:

Regularity conditions (C_R). There exist $\epsilon > 0$, a bivariate function r and a sequence ℓ_n such that, as $n \rightarrow \infty$,

$$(a) \quad \frac{\beta(\ell_n)}{\ell_n} n + \ell_n \frac{\log^2 k_n}{\sqrt{k_n}} \rightarrow 0;$$

$$(b) \quad \frac{n}{\ell_n k_n} Cov \left(\sum_{i=1}^{\ell_n} \mathbb{I}_{\{X_i > F^\leftarrow(1-k_n x/n)\}}, \sum_{i=1}^{\ell_n} \mathbb{I}_{\{X_i > F^\leftarrow(1-k_n y/n)\}} \right) \rightarrow r(x, y), \quad \forall 0 \leq x, y \leq 1 + \epsilon;$$

(c) For some constant C :

$$\frac{n}{\ell_n k_n} \mathbb{E} \left[\left(\sum_{i=1}^{\ell_n} \mathbb{I}_{\{F^\leftarrow(1-k_n y/n) < X_i \leq F^\leftarrow(1-k_n x/n)\}} \right)^4 \right] \leq C(y - x), \quad \forall 0 \leq x < y \leq 1 + \epsilon \text{ and } n \in \mathbb{N},$$

(5) established the asymptotic normality of $\hat{\gamma}_{k_n}^{(H)}$ as follows

$$\sqrt{k} (\hat{\gamma}_{k_n}^{(H)} - \gamma) \xrightarrow{d} \mathcal{N} \left(\frac{\lambda}{1 - \rho}, \sigma_H^2(\gamma) \right), \quad (1.3)$$

where $\sigma_H^2(\gamma) = \gamma^2 r(1, 1)$, with r the covariance structure, but has a simple expression in the i.i.d. context, where it is equal to γ^2 . In practice, the bias term of $\hat{\gamma}_{k_n}^{(H)}$ depends on whether ρ is close to zero or not, since under the second order condition (C_{SO}), the function $|A|$ is regularly varying at infinity with index ρ . This explains all the literature spread on bias correction in the i.i.d. context, see, e.g., (1), (8) and (11), etc. However, in the case of β -mixing time series, only the very recent papers by (2) and (16) deal with this problem and proposed a bias corrected estimator for γ . By the way, they established their asymptotic properties under the regularity conditions (C_R) and the second order assumptions (C_{SO}).

This paper is organised as follows. In Section 2, we introduce the estimation and the asymptotic properties of three adaptive unbiased tail index estimators and their corresponding high quantiles. We recall that these estimators come from two families of estimators, presented in (13), which generalize the Hill's estimator. And in Section 4, we give some examples of β -mixing series. Then in Section 5&6, we match our theoretical results with a simulation assessment and real data analysis in order to highlight the salients of our methods. Finally, Section 7 is devoted to the proofs of our main results.

2. Methodology and Main Results

In (13), two families of tail index estimators are introduced in the case of i.i.d. of extreme values. These estimators are parameterized by a positive real α and they also generalize the Hill's estimator. Using a Jackknife approach on these estimators, meaning asymptotically unbiased estimators for γ , the tail index have been constructed. In this work, we fit in our methodology with the case of β -mixing sequences to estimate a tail parameter such as the extreme value index or an extreme quantile.

2.1. Adaptive generalized tail index estimators. We consider the statistics introduced in (3) and defined by:

$$M_{k_n}^{(\alpha)} := \frac{1}{k_n} \sum_{i=1}^{k_n} (\log X_{n-i+1,n} - \log X_{n-k_n,n})^\alpha, \quad \alpha > 0.$$

This can be rewritten as a functional of $(Q_n(t) := X_{n-[k_n t],n})_{t \in [0,1]}$, and the tail quantile process as follows:

$$M_{k_n}^{(\alpha)} = \int_0^1 \left(\log \frac{Q_n(t)}{Q_n(1)} \right)^\alpha dt.$$

Now, to derive the asymptotic properties of this estimator, we first establish those of the quantile process and the moments $M_{k_n}^{(\alpha)}$. We show that the tail quantile process can be approximated by a Gaussian process as in the following results.

Proposition 2.1. *Suppose that (X_1, X_2, \dots) is a stationary β -mixing time series with continuous common marginal distribution function F and assume that (C_{SO}) and (C_R) hold. Let k_n be an intermediate sequence satisfying $k_n \rightarrow \infty$, $k_n/n \rightarrow 0$ and $k_n^{1/2}A(n/k_n) = O(1)$, as $n \rightarrow \infty$. Then, for given α , $\delta > 0$, under a Skorohod construction, there exist a function $\tilde{A} \sim A$, and a centered Gaussian process $(W(t))_{t \in [0,1]}$ with covariance function r , such that, as $n \rightarrow \infty$:*

$$\sup_{t \in (0,1]} t^{1/2+\delta} \left| \sqrt{k} \left(\log \frac{Q_n(t)}{Q_n(1)} \right)^\alpha - \gamma^\alpha (-\log t)^\alpha - \alpha \gamma^\alpha (-\log t)^{\alpha-1} (t^{-1}W(t) - W(1)) - \sqrt{k} \tilde{A}(n/k) \alpha \gamma^{\alpha-1} (-\log t)^{\alpha-1} \frac{t^{-\rho} - 1}{\rho} \right| \rightarrow 0 \text{ a.s.}$$

The aim of the following corollary is to provide the asymptotic properties of the moments $M_{k_n}^{(\alpha)}$ by applying the Proposition 2.1.

Corollary 2.1. *Assume that the conditions in the Proposition 2.1 hold. Then under the same Skorohod construction as in Proposition 2.1, as $n \rightarrow \infty$,*

$$\sqrt{k_n} \left(M_{k_n}^{(\alpha)} - \gamma^\alpha \Gamma(\alpha + 1) \right) - \alpha \gamma^\alpha P^{(\alpha)} - \sqrt{k_n} \tilde{A}(n/k_n) \gamma^{\alpha-1} \frac{\Gamma(\alpha + 1)}{\rho} \left(\frac{1}{(1-\rho)^\alpha} - 1 \right) \rightarrow 0 \text{ a.s.,}$$

where the process term

$$P^{(\alpha)} = \int_0^1 (-\log t)^{\alpha-1} (t^{-1}W(t) - W(1)) dt$$

is a normally distributed random variable with mean zero and covariance $\text{Cov}(P^{(\alpha)}, P^{(\beta)}) = c_{\alpha, \beta}$ defined as,

$$c_{\alpha, \beta} = \int_0^1 \int_0^1 (-\log s)^{\alpha-1} (-\log t)^{\beta-1} \times \left(\frac{r(s, t)}{st} - \frac{r(s, 1)}{s} - \frac{r(1, t)}{t} + r(1, 1) \right) ds dt,$$

with the covariance structure r defined as in the regularity condition (C_R) .

The case $\alpha = 1$ is special since its corresponding tail index estimator $M_{k_n}^{(1)}$ is the Hill's estimator. For this reason, the asymptotic variance $\sigma_H^2(\gamma)$ is equal to $\gamma^2 c_{1,1}$, with $c_{1,1} = r(1, 1)$ (see, eg. (5)). Following the idea in the Corollary 2.1 and the methodology in (13), we can define two families of tail index estimators which generalize the Hill's estimator. They are also parameterized by a positive real α . The first one is defined by:

$$\tilde{\gamma}_{k_n}^{(\alpha)} = \left(\frac{M_{k_n}^{(\alpha)}}{\Gamma(\alpha + 1)} \right)^{1/\alpha}, \quad \alpha > 0 \quad (2.4)$$

and the second one by:

$$\hat{\gamma}_{k_n}^{(\alpha)} = \frac{M_{k_n}^{(\alpha)}}{\Gamma(\alpha + 1) \left(M_{k_n}^{(1)} \right)^{\alpha-1}}, \quad \alpha \geq 1. \quad (2.5)$$

It is clear that the Hill's estimator $\hat{\gamma}_{k_n}^{(H)}$ corresponds to $M_{k_n}^{(1)} = \tilde{\gamma}_{k_n}^{(1)} = \hat{\gamma}_{k_n}^{(1)}$. The following theorem gives the asymptotic expansions of the estimators $\tilde{\gamma}_{k_n}^{(\alpha)}$ and $\hat{\gamma}_{k_n}^{(\alpha)}$ in terms of Gaussian process $P^{(\alpha)}$.

Theorem 2.1. *Assume that the conditions in the Proposition 2.1 hold. Then we have, as $n \rightarrow \infty$:*

$$\tilde{\gamma}_{k_n}^{(\alpha)} \stackrel{d}{=} \gamma + \frac{\gamma P^{(\alpha)}}{k_n^{1/2} \Gamma(\alpha + 1)} + \tilde{A}(n/k_n) \frac{1 - (1 - \rho)^\alpha}{\alpha \rho (1 - \rho)^\alpha} + o_{\mathbb{P}}(k_n^{-1/2}) \quad \alpha > 0 \quad (2.6)$$

and for $\alpha \geq 1$

$$\hat{\gamma}_{k_n}^{(\alpha)} \stackrel{d}{=} \gamma + \frac{\gamma \alpha P^{(\alpha)}}{k_n^{1/2} \Gamma(\alpha + 1)} - \gamma(\alpha - 1) \frac{P^{(1)}}{k_n^{1/2}} + \tilde{A}(n/k_n) \left\{ \frac{1 - (1 - \rho)^\alpha}{\rho(1 - \rho)^\alpha} - \frac{\alpha - 1}{1 - \rho} \right\} + o_{\mathbb{P}}(k_n^{-1/2}). \quad (2.7)$$

The following corollary establishes the asymptotic normality of the generalized tail index estimators $\tilde{\gamma}_{k_n}^{(\alpha)}$ and $\hat{\gamma}_{k_n}^{(\alpha)}$.

Corollary 2.2. *Assume that the conditions in the Proposition 2.1 hold. Suppose that k_n is an intermediate sequence satisfying $k_n \rightarrow \infty$, $k_n/n \rightarrow 0$ and $\sqrt{k_n} A(n/k_n) \rightarrow \lambda \in \mathbb{R}$, as $n \rightarrow \infty$. Then we have, as $n \rightarrow \infty$:*

$$\sqrt{k_n} (\tilde{\gamma}_{k_n}^{(\alpha)} - \gamma) \xrightarrow{d} \mathcal{N} \left(\lambda \frac{1 - (1 - \rho)^\alpha}{\alpha \rho (1 - \rho)^\alpha}, \frac{\gamma^2 c_{\alpha, \alpha}}{\Gamma^2(\alpha + 1)} \right)$$

and

$$\sqrt{k_n} (\hat{\gamma}_{k_n}^{(\alpha)} - \gamma) \xrightarrow{d} \mathcal{N} \left(\lambda \left\{ \frac{1 - (1 - \rho)^\alpha}{\rho(1 - \rho)^\alpha} - \frac{\alpha - 1}{1 - \rho} \right\}, \gamma^2 \left\{ (\alpha - 1)^2 c_{1,1} + \frac{c_{\alpha, \alpha}}{\Gamma^2(\alpha + 1)} - 2 \frac{(\alpha - 1) c_{1, \alpha}}{\Gamma(\alpha + 1)} \right\} \right),$$

where $c_{1,1}$, $c_{1,\alpha}$ and $c_{\alpha,\alpha}$ are as defined in the Corollary 2.1.

Remark 2.1. From the Theorem 2.1, the results show that the bias terms in the i.i.d. case (13) remain unchanged in the case of β -mixing. This can be explained by the fact that the stationary β -mixing data satisfying the regularity condition (C_R) depend on the covariance function r which impact the variance terms of the families of estimators and not their bias terms. In addition, the analysis of (13) show that the bias terms are minimal for some suitable values of (ρ, α) .

2.2. Asymptotically unbiased estimators of the tail index . From the Corollary 2.2, it is clear that the second order parameter ρ is important for the reduction of the bias of the tail index estimators. In the case of i.i.d. random variables, the bias reduction problem have received much attention from many authors such as (4), (9),(10), (14), (23), and (28). Nevertheless, only few authors considered the serial dependence topic. We can mention (5), (6) and (18), and very recently (2) and (16) as references.

The purpose in this section is to introduce a class of reduced bias estimators by using a jackknife approach. Following the idea of (14) and from the Remark 2.2, let's choose the values 1 and 2 for α . The resulting tail index estimators are:

$$\widehat{\gamma}_{k_n}^{(H)} = \widetilde{\gamma}_{k_n}^{(1)} = \widehat{\gamma}_{k_n}^{(1)} = M_{k_n}^{(1)}, \quad \widehat{\gamma}_{k_n}^{(2)} = \frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} \quad \text{and} \quad \widetilde{\gamma}_{k_n}^{(2)} = \sqrt{\frac{M_{k_n}^{(2)}}{2}}. \quad (2.8)$$

From any pair of these above mentioned statistics, we construct the following adaptive reduced bias tail index estimators in the generalized Jackknife methodology sense (See eg, (14)):

$$\begin{aligned} \widehat{\gamma}_{k_n, \widehat{\rho}}^{(Ub1)} &:= \frac{1}{\widehat{\rho}} \left(\sqrt{2M_{k_n}^{(2)}} - (2 - \widehat{\rho}) \frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} \right), \\ \widehat{\gamma}_{k_n, \widehat{\rho}}^{(Ub2)} &:= \frac{1}{\widehat{\rho}} \left((2 - \widehat{\rho})M_{k_n}^{(1)} - (1 - \widehat{\rho})\sqrt{2M_{k_n}^{(2)}} \right) \end{aligned}$$

and

$$\widehat{\gamma}_{k_n, \widehat{\rho}}^{(dH)} := \frac{1}{\widehat{\rho}} \left(M_{k_n}^{(1)} - (1 - \widehat{\rho}) \frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} \right),$$

where $\widehat{\rho}$ is either a canonical negative value $\widehat{\rho} := \rho = \rho_0$ or a consistent estimator $\widehat{\rho} := \widehat{\rho}_{k_\rho}$ of ρ , with $k_\rho := k_{\rho, n}$ an intermediate sequence of integers greater than k_n , satisfying $k_\rho \rightarrow \infty$ and $k_\rho/n \rightarrow 0$, as $n \rightarrow \infty$. Note that the estimator $\widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(dH)}$ can be rewritten as

$$\widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(dH)} = M_{k_n}^{(1)} - \frac{M_{k_n}^{(2)} - 2 \left(M_{k_n}^{(1)} \right)^2}{2M_{k_n}^{(1)} \widehat{\rho}_{k_\rho} \left(1 - \widehat{\rho}_{k_\rho} \right)^{-1}}.$$

Clearly, this estimator is seen as a corrected bias version of the Hill's estimator $\widehat{\gamma}_{k_n}^{(H)}$. It has been studied in the i.i.d. case by (23). (19) showed that for the canonical choice $\rho_0 = -1$, the estimator $\widehat{\gamma}_{k_n, \rho_0}^{(dH)}$ is the t-Hill log-gamma (t-lgHill) estimator of the tail index γ . They established its weak consistency for the moving average process. Recently, (2) adapted the estimator $\widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(dH)}$ in the β -mixing serials case and established its asymptotic normality related to the β -mixing regularity conditions and a third order assumption.

The purpose of the following theorem is to establish only the asymptotic normality of the estimators $\widehat{\gamma}_{k_n, \widehat{\rho}}^{(Ub1)}$, $\widehat{\gamma}_{k_n, \widehat{\rho}}^{(Ub2)}$ and $\widehat{\gamma}_{k_n, \widehat{\rho}}^{(dH)}$ under the β -mixing regularity conditions and the second order assumption which is less strong than the third order condition (see (2)).

Theorem 2.2. *Let (X_1, X_2, \dots) be a stationary β -mixing time series with a continuous common marginal distribution function F and assume that (C_{SO}) and (C_R) hold. Let $\widehat{\rho}$ be either a canonical negative value $\widehat{\rho} := \rho = \rho_0$ or an estimator $\widehat{\rho} := \widehat{\rho}_{k_\rho}$ of ρ , consistent in probability such that $k_\rho := k_{\rho, n}$ is an intermediate sequence of integers satisfying $k_\rho \rightarrow \infty$ and $k_\rho/n \rightarrow 0$, as $n \rightarrow \infty$. If k_n is another intermediate sequence such that $k_n \rightarrow \infty$, $k_n/n \rightarrow 0$ and $\sqrt{k_n}A(n/k_n) \rightarrow \lambda \in \mathbb{R}$, as $n \rightarrow \infty$, then we have:*

$$\begin{aligned} \sqrt{k_n} \left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(Ub1)} - \gamma \right) &\xrightarrow{d} \mathcal{N} \left(0, \sigma^2(\gamma, \rho) \right), \\ \sqrt{k_n} \left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(Ub2)} - \gamma \right) &\xrightarrow{d} \mathcal{N} \left(0, \sigma^2(\gamma, \rho) \right) \end{aligned}$$

and

$$\sqrt{k_n} \left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(dH)} - \gamma \right) \xrightarrow{d} \mathcal{N} \left(0, \sigma^2(\gamma, \rho) \right),$$

where

$$\sigma^2(\gamma, \rho) = \frac{\gamma^2}{\rho} \left((2 - \rho)^2 c_{1,1} + (1 - \rho)^2 c_{2,2} + 2(2 - \rho)(-1)c_{1,2} \right).$$

Remark 2.2. *The Theorem 4.1 in (2) states a similar result for the estimator $\widehat{\gamma}_{k_n}^{(dH)}(\widehat{\rho}_{k_\rho})$. Comparing this theorem to the assumptions of our Theorem 2.2, there are less constraints, particularly, we do not need a third order condition and only $\sqrt{k_\rho}A(n/k_\rho) \rightarrow \infty$, as $n \rightarrow \infty$ is needed on the intermediate sequence k_ρ . This is because we only require the consistency in probability for the estimator $\widehat{\rho}_{k_\rho}$, not its asymptotic normality. However, our rate of convergence $\sqrt{k_n}$ is smaller than the one obtained in (2) since the assumption $\sqrt{k_n}A(n/k_n) \rightarrow \infty$, as $n \rightarrow \infty$, was used instead of our condition, $\sqrt{k_n}A(n/k_n) \rightarrow \lambda \in \mathbb{R}$, as $n \rightarrow \infty$. Though, their rate of convergence has the form $\sqrt{k_n}$, their intermediate sequence is larger than ours. This occurs when $|A|$, the rate function, varies regularly at infinity with index ρ .*

A possible choice for $\widehat{\rho}_{k_\rho}$ is the most performed estimator among those studied in the i.i.d. case (see, e.g, (4), (14)) and used in the β -mixing case by (2; 16):

$$\widehat{\rho}_{k_\rho}^{(*)} = \frac{6S_{k_\rho}^{(2)} - 4 + \sqrt{3S_{k_\rho}^{(2)} - 2}}{4S_{k_\rho}^{(2)} - 3}, \quad \text{provided } S_{k_\rho}^{(2)} \in \left(\frac{2}{3}, \frac{3}{4} \right), \quad (2.9)$$

where

$$S_{k_\rho}^{(2)} = \frac{3 \left[M_{k_\rho}^{(4)} - 24 \left(M_{k_\rho}^{(1)} \right)^4 \right] \left[M_{k_\rho}^{(2)} - 2 \left(M_{k_\rho}^{(1)} \right)^2 \right]}{4 \left[M_{k_\rho}^{(3)} - 6 \left(M_{k_\rho}^{(1)} \right)^3 \right]^2}.$$

The asymptotic properties of $\widehat{\rho}_{k_\rho}^{(*)}$ have been established in (2) in the case of β -mixing serials.

Since the mentioned jackknife estimators have similar asymptotic variances, they can be denoted by $\widehat{\gamma}_{k_n, \hat{\rho}}^{(\bullet)}$. From the previous remark, follow the direct consequences in the corollary:

Corollary 2.3. *Let (X_1, X_2, \dots) be a stationary β -mixing time series with a continuous common marginal distribution function F and assume that (C_{SO}) and (C_R) hold. Let $\widehat{\rho}_{k_\rho}^{(*)}$ be the estimator for ρ defined in (2.9), where the intermediate sequence $k_\rho := k_{\rho, n}$ satisfies $k_\rho \rightarrow \infty$, $k_\rho/n \rightarrow 0$ and $\sqrt{k_\rho}A(n/k_\rho) \rightarrow \infty$, as $n \rightarrow \infty$. If k_n is another intermediate sequence such that $k_n \rightarrow \infty$, $k_n/n \rightarrow 0$ and $\sqrt{k_n}A(n/k_n) \rightarrow \lambda \in \mathbb{R}$, as $n \rightarrow \infty$, then we have:*

$$\sqrt{k_n}(\widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}^{(*)}}^{(\bullet)} - \gamma) \xrightarrow{d} \mathcal{N}(0, \sigma^2(\gamma, \rho)),$$

where $\sigma(\gamma, \rho)$ is given in Theorem 2.2.

2.3. Estimation of extreme quantiles. The quantile, at probability level $(1-t) \in (0, 1)$ with respect to F denoted by $x(t)$, is defined as:

$$x(t) := U(1/t). \quad (2.10)$$

Therefore $x(t)$, the quantile is estimated by $\widehat{x}(t) := X_{n-[nt]; n}$. Consider some positive sequence $t = t_n$ which tends to 0, as $n \rightarrow \infty$. Then it is possible to establish the consistency of $\widehat{x}(t_n)$, the non parametric estimator when $t_n \rightarrow 0$ slowly enough. However, in some fields in life such as insurance, finance, hydrology and reliability, a major requirement is to find values, large enough so that the chances of exceeding them are very small. This leads to remove the restriction on the rate of convergence of t_n to 0. Moreover, the interest is to estimate $x(p)$, an extreme quantile, where p , the tail probability depends on the observed sample size n (i.e $p := p_n$) and p_n is smaller than $1/n$. Hence, it is not possible to have a non parametric estimate of such a quantile.

The goal of this section is to tackle this estimation problem in a β -mixing serials framework in order to estimate $x(p) = U(1/p)$, the extreme quantile with $np < 1$. Following (21; 2), we propose, in a biased reduction method, to estimate $x(p)$, the extreme quantile. The construction of our bias reduction procedure is based on our second-order condition (C_{SO}) which states

$$\frac{U(tx)}{U(t)} \approx x^\gamma \left\{ 1 - \rho^{-1} A(t)[1 - x^\rho] \right\}, \quad t \rightarrow \infty.$$

Let $tx = 1/p$ and $t = n/k_n \rightarrow \infty$, as $n \rightarrow \infty$. We obtain the following approximation:

$$x(p) = U(1/p) \approx U(n/k_n) \left(\frac{k_n}{np} \right)^\gamma \left\{ 1 - \rho^{-1} A(n/k_n) \left[1 - \left(\frac{k_n}{np} \right)^\rho \right] \right\}, \quad (2.11)$$

where γ , ρ and $A(n/k_n)$ are unknown. The first part $U(n/k_n) (k/(np))^\gamma$ in the right side of (2.11) is exactly estimated by the Weissman's estimator ((27)) $\widehat{x}_{k_n}^{(W)}(p)$ and defined as:

$$\widehat{x}_{k_n}^{(W)}(p) = X_{n-k_n, n} \left(\frac{k_n}{np} \right)^{\widehat{\gamma}_{k_n}^{(H)}},$$

where $X_{n-k_n, n}$ is the empirical estimator of $U(n/k_n)$ and $\hat{\gamma}_{k_n}^{(H)}$ is the Hill's estimator of γ . Obviously, $\hat{x}_{k_n}^{(W)}(p)$, the Weissman's estimator exhibits a potential bias because it depends on the Hill's estimator $\hat{\gamma}_{k_n}^{(H)}$ which has a similar problem. The expression $1 - \rho^{-1}A(n/k_n)[1 - (k_n/(np))^\rho]$ can be viewed as a correcting term since $A(n/k_n)$ tends to 0. And this leads to the need to estimate $A(n/k_n)$ and ρ .

Using the tail index estimators in (2.8) and their expansions in Theorem 2.1, we obtain

$$\frac{\frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} - M_{k_n}^{(1)}}{A(n/k_n)} \xrightarrow{\mathbb{P}} \frac{\rho}{(1-\rho)^2}. \text{ As } n \rightarrow \infty \quad (2.12)$$

From equation 2.12, by substituting ρ with $\hat{\rho}$, we estimate $A(n/k_n)$ as

$$\hat{A}(n/k_n) = \frac{(1-\hat{\rho})^2}{\hat{\rho}} \left(\frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} - M_{k_n}^{(1)} \right).$$

Finally, the use of the estimators and the approximation in Section 2.2 and (2.11) respectively lead to the following class of reduced bias estimators of $x(p)$, the extreme quantile:

$$\begin{pmatrix} \hat{x}_{k_n, \hat{\rho}}^{(Ub1)}(p) \\ \hat{x}_{k_n, \hat{\rho}}^{(Ub2)}(p) \\ \hat{x}_{k_n, \hat{\rho}}^{(dH)}(p) \end{pmatrix} = X_{n-k_n, n} \times \begin{pmatrix} \left(\frac{k}{np} \right)^{\hat{\gamma}_{k_n, \hat{\rho}}^{(Ub1)}} \\ \left(\frac{k}{np} \right)^{\hat{\gamma}_{k_n, \hat{\rho}}^{(Ub2)}} \\ \left(\frac{k}{np} \right)^{\hat{\gamma}_{k_n, \hat{\rho}}^{(dH)}} \end{pmatrix} \times \left\{ 1 - \frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)} \hat{\rho}^2 (1-\hat{\rho})^{-2}} \left[1 - \left(\frac{k}{np} \right)^{\hat{\rho}_{k_\rho}} \right] \right\}. \quad (2.13)$$

The asymptotic normality of extreme quantile estimators is established in the following theorem.

Theorem 2.3. *Let (X_1, X_2, \dots) be a stationary β -mixing time series with a continuous common marginal distribution function F and assume that (C_{SO}) and (C_R) hold. Let $\hat{\rho}$ be either a canonical negative value $\hat{\rho} := \rho = \rho_0$ or an estimator $\hat{\rho} := \hat{\rho}_{k_\rho}$ of ρ , consistent in probability such that $k_\rho := k_{\rho, n}$ is an intermediate sequence of integers satisfying $k_\rho \rightarrow \infty$ and $k_\rho/n \rightarrow 0$, as $n \rightarrow \infty$. Consider another intermediate sequence such that $n \rightarrow \infty, np_n/k_n \rightarrow 0$ and $\log(np_n)/\sqrt{k_n} \rightarrow 0$. Then*

$$\frac{\sqrt{k_n}}{\log(np_n)} \left(\frac{\hat{x}_{k_n, \hat{\rho}}^{(\bullet)}(p)}{x(p)} - 1 \right) \rightarrow \mathcal{N}(0, \sigma^2(\gamma, \rho)),$$

where $\sigma^2(\gamma, \rho)$ is defined in Theorem 2.2.

3. EXAMPLES OF β -MIXING SERIES

In this section, we present some examples dealing with classical stationary models satisfying the regularity (C_R) assumptions.

Example 3.1. (*Autoregressive (AR) model*)

Consider the stationary solution of AR(1) equation:

$$X_i = \theta X_{i-1} + \varepsilon_i, \quad (3.14)$$

for some $\theta \in (0, 1)$ and i.i.d. random variables ε_i . The distribution function of the innovations is denoted by F_ε . Assume that F_ε admits a positive Lebesgue density which is L_1 -Lipschitz-continuous; see (6) eq. (4.2). Suppose that as $x \rightarrow \infty$,

$1 - F_\varepsilon(x) \sim px^{-1/\gamma}l(x)$ and $F_\varepsilon(-x) \sim qx^{-1/\gamma}l(x)$, for some slowly varying function l and $p = 1 - q \in (0, 1)$.

Then from Sect. 3.2 of (6), we obtain $1 - F(x) \sim d_\theta(1 - F(x))$ as $x \rightarrow \infty$, where $d_\theta = (1 - \theta^{1/\gamma})^{-1}$.

Furthermore, the regularity conditions hold with,

$$r(x, y) = x \wedge y + \sum_{m=1}^{\infty} (c_m(x, y) + c_m(y, x));$$

where $c_m(x, y) = x \wedge y \theta^{m/\gamma}$.

Example 3.2. (*Moving average (MA) model*)

Consider the stationary solution of MA(1) equation:

$$X_i = \theta \varepsilon_{i-1} + \varepsilon_i, \quad (3.15)$$

where the innovation ε satisfies the same conditions as in the AR(1) model (example 4.1). And from Sect. 3.2 of (6), we obtain $1 - F(x) \sim d_\theta(1 - F_\varepsilon(x))$ as $x \rightarrow \infty$, where $d_\theta = 1 + \theta^{1/\gamma}$. One can also compute $r(x, y) = x \wedge y + (1 + \theta^{1/\gamma})^{-1}(x \wedge y \theta^{1/\gamma} + y \wedge x \theta^{1/\gamma})$.

Example 3.3. (*Generalized autoregressive conditional heteroskedasticity (GARCH) model*)

Consider the stationary solution to the recursive system of equations:

$$\begin{cases} X_t = & \sigma_t \varepsilon_t, \\ \sigma_t^2 = & \lambda_0 + \lambda_1 X_{t-1}^2 + \lambda_2 \sigma_{t-1}^2, \end{cases} \quad (3.16)$$

where the process of innovations ε are i.i.d. with zero mean and unit variance. The stationary solution X_t of this GARCH(1,1) model follows a heavy-tailed distribution, even if the innovations ε_t are normally distributed, see (20) and (7).

This GARCH(1,1) model satisfies the regularity conditions (C_R) (see (25) and (5)), but without an explicit covariance function $r(\cdot, \cdot)$.

4. Examples of β -mixing series

In this section, we present some examples dealing with classical stationary models satisfying the regularity (C_R) assumptions.

Example 4.1. (*Autoregressive (AR) model*)

Consider the stationary solution of AR(1) equation:

$$X_i = \theta X_{i-1} + \varepsilon_i, \quad (4.17)$$

for some $\theta \in (0, 1)$ and i.i.d. random variables ε_i . The distribution function of the innovations is denoted by F_ε . Assume that F_ε admits a positive Lebesgue density which is L_1 -Lipschitz-continuous; see (6) eq. (42). Suppose that as $x \rightarrow \infty$,

$1 - F_\varepsilon(x) \sim px^{-1/\gamma}l(x)$ and $F_\varepsilon(-x) \sim qx^{-1/\gamma}l(x)$, for some slowly varying function l and $p = 1 - q \in (0, 1)$. Then from Sect. 3.2 of (6), we obtain $1 - F(x) \sim d_\theta(1 - F_\varepsilon(x))$ as $x \rightarrow \infty$, where $d_\theta = (1 - \theta^{1/\gamma})^{-1}$. Furthermore, the regularity conditions hold with,

$$r(x, y) = x \wedge y + \sum_{m=1}^{\infty} (c_m(x, y) + c_m(y, x));$$

where $c_m(x, y) = x \wedge y \theta^{m/\gamma}$.

Example 4.2. (Moving average (MA) model)

Consider the stationary solution of MA(1) equation:

$$X_i = \theta \varepsilon_{i-1} + \varepsilon_i, \quad (4.18)$$

where the innovation ε satisfies the same conditions as in the AR(1) model (example 4.1). And from Sect. 3.2 of (6), we obtain $1 - F(x) \sim d_\theta(1 - F_\varepsilon(x))$ as $x \rightarrow \infty$, where $d_\theta = 1 + \theta^{1/\gamma}$. One can also compute $r(x, y) = x \wedge y + (1 + \theta^{1/\gamma})^{-1}(x \wedge y \theta^{1/\gamma} + y \wedge x \theta^{1/\gamma})$.

Example 4.3. (Generalized autoregressive conditional heteroskedasticity (GARCH) model)

Consider the stationary solution to the recursive system of equations:

$$\begin{cases} X_t = & \sigma_t \varepsilon_t, \\ \sigma_t^2 = & \lambda_0 + \lambda_1 X_{t-1}^2 + \lambda_2 \sigma_{t-1}^2, \end{cases} \quad (4.19)$$

where the process of innovations ε are i.i.d. with zero mean and unit variance. The stationary solution X_t of this GARCH(1,1) model follows a heavy-tailed distribution, even if the innovations ε_t are normally distributed, see (20) and (7).

This GARCH(1,1) model satisfies the regularity conditions (C_R) (see (25) and (5)), but without an explicit covariance function $r(., .)$.

5. Simulation study

In this section, we proceed by generating the data for the four (04) models. This involves an independent model and the three models mentioned above, Examples 4.1 - 4.3. We define ε such that:

$$F_\varepsilon(\varepsilon) = \begin{cases} (1 - q)(1 - \tilde{F}(-\varepsilon)) & \text{if } \varepsilon < 0, \\ 1 - q + q\tilde{F}(\varepsilon) & \text{if } \varepsilon > 0, \end{cases}$$

where \tilde{F} stands for the unit Fréchet distribution function. For $q = 0.75$, F_ε belongs to the max-domain of attraction with $\gamma = 1$, an extreme value index. We generate $N = 1000$ time series of size $n = 1000$ based on i.i.d. observations from F_ε . This concerns all four models. The theoretical value of $\gamma = 1$ concerns the three first models. The fourth one is 0.258:

- **Model 1.** Independence model with $X_t = \varepsilon_t$ (as a particular case of AR(1) with $\theta = 0$). The theoretical value of $x(0.001)$ is 749.80.
- **Model 2.** AR(1) model (4.17) with $\theta = 0.3$. The theoretical value of $x(0.001)$ is 1072.26.
- **Model 3.** MA(1) model (4.18) with $\theta = 0.3$. The theoretical value of $x(0.001)$ is 972.85.
- **Model 4.** GARCH(1,1) model (6.20) with standardized Student t innovations with 5.64 degrees of freedom and $\lambda_0 = 8.26 \times 10^{-07}$, $\lambda_1 = 0.052$, $\lambda_2 = 0.941$. The theoretical value of $x(0.001)$ is 0.0592.

Note that the theoretical values are computed by the Monte Carlo method based on 1000 samples of size 10^6 , see (2). For each model, we generate $N = 1000$ samples of size $n = 1000$. Firstly, we focus on the extreme value index γ . We apply the four (04) tail index estimators: $\hat{\gamma}_{k_n}^{(H)}$, $\hat{\gamma}_{k_n, k_\rho}^{(Ub1)}$, $\hat{\gamma}_{k_n, k_\rho}^{(Ub2)}$ and $\hat{\gamma}_{k_n, k_\rho}^{(dH)}$, with their associated extreme quantile estimators $\hat{x}_{k_n}^{(W)}(p)$, $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}(p)$, $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}(p)$ and $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(dH)}(p)$, with $p = 0.001$, on each model in order to compare them. But for the estimator $\hat{\rho}_{k_\rho}$, we use $\hat{\rho}_{k_\rho}^{(*)}$ defined in (2.9), where the sequence k_ρ is selected as follows:

$$k_\rho := \sup \left\{ k : k \leq \min \left(m - 1, \frac{2m}{\log \log m} \right) \text{ and } \hat{\rho}_k \text{ exists} \right\},$$

where m is the number of positive observations in the sample.

Secondly, on the one hand we compare the performance of the tail index estimators in each model. And on the other, we do the same for the extreme quantiles estimators. For this reason, we compute the absolute value of the mean of the bias (ABias) together with the root mean squared errors (RMSE) based on the N samples, and defined as

$$\text{ABias}(\eta, k_n) := \left| \frac{1}{N} \sum_{i=1}^N \frac{\hat{\eta}^{(i)}}{\eta} - 1 \right| \quad \text{and} \quad \text{RMSE}(\eta, k) := \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{\hat{\eta}^{(i)}}{\eta} - 1 \right)^2},$$

where η is either γ or $x(p)$, and $\hat{\eta}^{(i)}$ is the i -th value ($i = 1, \dots, N$) of an estimator of γ or $x(p)$ evaluated at k_n .

In Figure 5.1, resp. Figure 5.2, we plot the results against the corresponding k_n values of the tail index estimators, resp. the extreme quantile estimators, for each of the models by row, and by column, the ABias (left) and RMSE (right). We have four curves in each graph. The full line corresponds to $\hat{\gamma}_{k_n}^{(H)}$ (resp. $\hat{x}_{k_n}^{(W)}(p)$) estimator, the dashed line to $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(dH)}$ (resp. $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(dH)}(p)$) estimator, the dotted line to $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}$ (resp. $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}(p)$) estimator and the longdashed line to $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}$ (resp. $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}(p)$) estimator.

By these simulations, we observe that:

- (1) The Figure 5.1 shows that our tail index estimators ($\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}$ and $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}$) look the same as the $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(dH)}$ estimator given in (2), in term of ABias and RMSE for the different Models. In Models 1, 2 and 3 especially, we observe a longer stability of the three asymptotically unbiased estimators as a function of k_n , which is not the case for the Hill's estimator $\hat{\gamma}_{k_n}^{(H)}$. Our first estimator $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}$ sticks at a lower level for that. However, the Hill's estimator shows better performance for small values of k_n . Now, regarding the *GARCH*(1, 1) model, the Hill's estimator performs very poorly, whereas our two unbiased estimators $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}$ and $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}$ are at least as good as $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(dH)}$ estimator with the best performance for $\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}$ estimator.
- (2) In Figure 5.1, our extreme quantile estimators ($\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}(p)$ and $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}(p)$) follow the same profile as $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(dH)}(p)$ (introduced in (2)), in term of ABias and RMSE for the different Models with a longer stability as a function of k_n , better than the Weissman's estimator ($\hat{x}_{k_n}^{(W)}(p)$). In Models 1, 2 and 3, we observe that our unbiased estimator $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}(p)$ remains at a lower level. However, the Weissman's estimator shows better performance for small values of k_n . In Model 4, our unbiased estimator $\hat{x}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}(p)$ turns out best compared to the others.

In summary, this simulation study points out that the Generalized Jackknife methodology which consists of removing the bias, the extreme value index and extreme quantile estimators remain stable for a broad range of k_n values even if the dataset exhibits serial dependence. Therefore, the bias reduction procedure related to the serial dependence case enables to overcome the problem of the choice of the intermediate sequence k_n used in the application of extreme value statistics to financial time series.

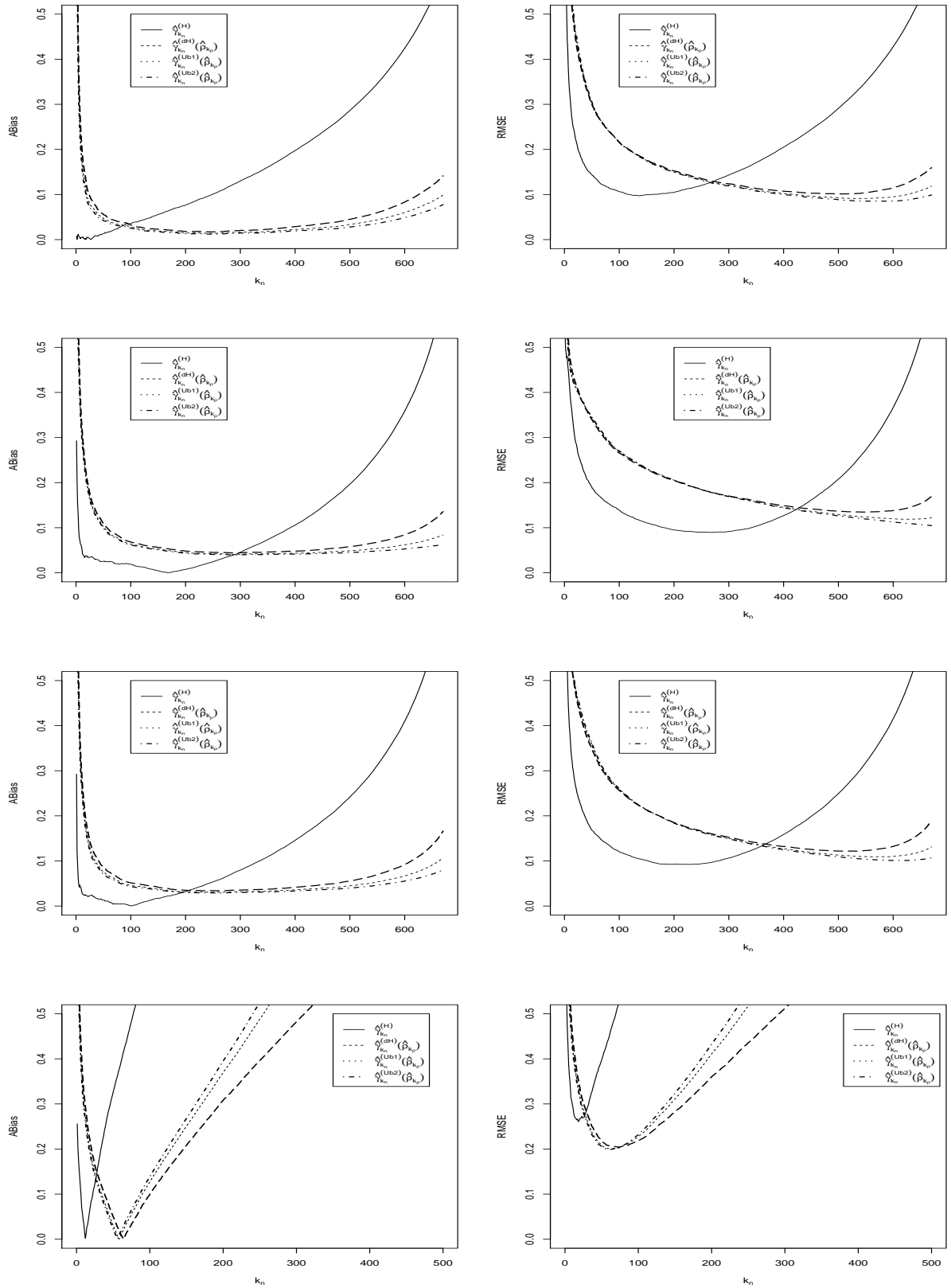


FIGURE 5.1. Simulation of the tail index: By row, Models 1, 2, 3, 4. By column, ABias (*left*) and RMSE (*right*) as functions of k_n .

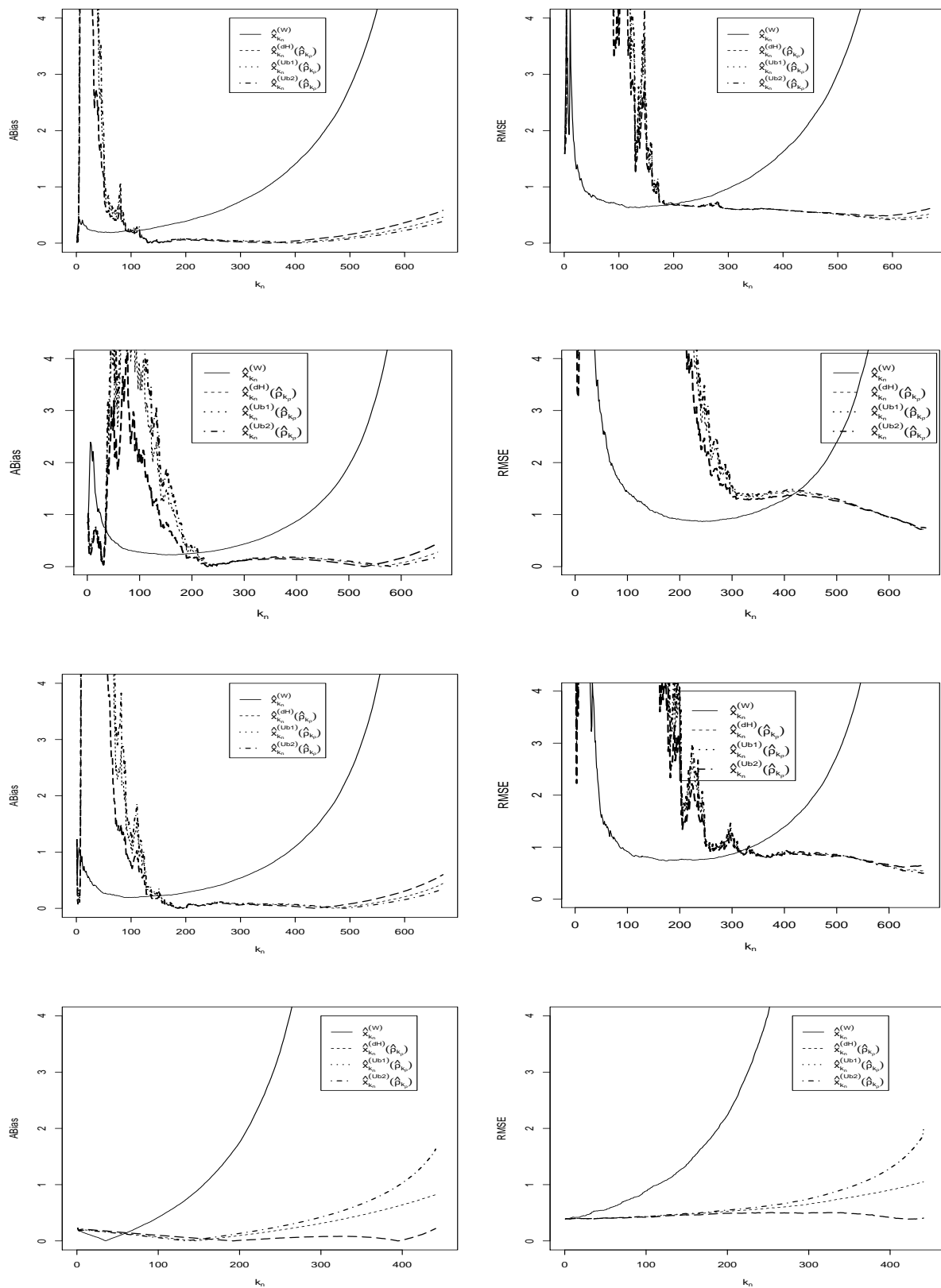


FIGURE 5.2. Simulation of the extreme quantiles: By row, **Models 1, 2, 3, 4**. By column, ABias (*left*) and RMSE (*right*) as functions of k_n .

6. APPLICATION TO FINANCIAL MARKET INDEX

In this section, we illustrate the performance of our bias reduction procedure through a real study case dealing with the S&P500 index financial data. We use the tail index and extreme quantiles estimators to assess the S&P500 index: I_t , $t = 1, \dots, n$. Figure 6.3 shows the daily negative log-returns $X_t = \log(I_t/I_{t-1}) < 0$, (*Loss*) for $n = 1000$ values of S&P500 index from *April 23th*, 2018 to *April 8th*, 2022. In a risk management perspective, the Value-at-Risk (VaR) is a common quantity hardwired in the international regulatory framework referred to as the Basel Accords (see (26), for a historical review of the Basel International Settlement). The Basel Accord requires the largest international banks to hold regulatory capital for the trading book based on a 99%-VaR over a 1-day or 10-day holding period. The VaR-based risk capital calculation drew a lot of attention over the last two decades (see (22), Chapter 1). The α -VaR for the horizon $h = 1$ day is the quantile $x(p) = U(1/p)$ with $p = 1 - \alpha$ of the distribution for the index daily log-returns. From Figure 6.3, we observe that although the loss return series can be regarded

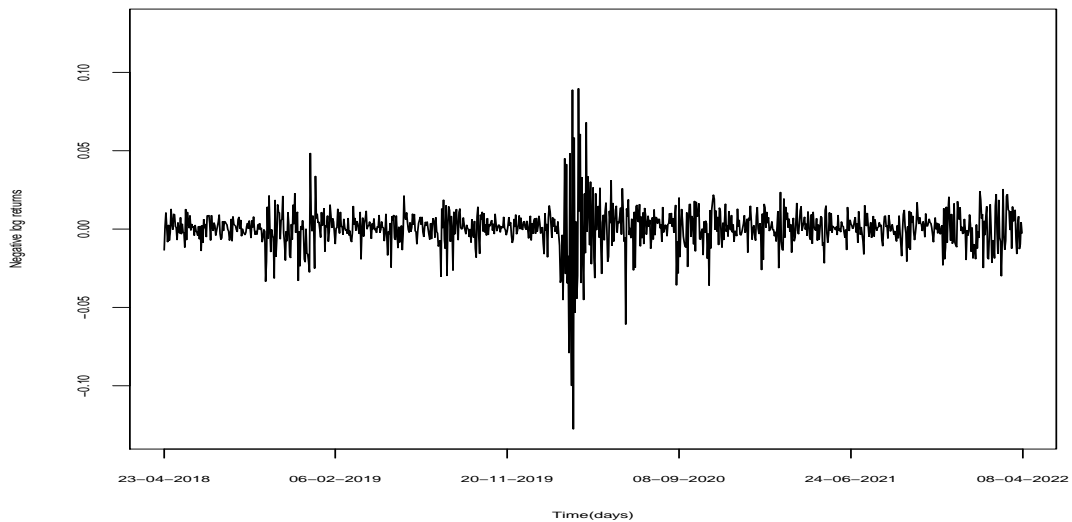


FIGURE 6.3. S&P500 index data: daily negative log-returns from *April 23th*, 2018 to *April 8th*, 2022.

as stationary, there is evidence of serial dependence such as cluster of volatility. To model this kind of phenomenon, the following $GARCH(1,1)$ model (6.20) is appropriate :

$$\begin{cases} X_t = & \sigma_t \varepsilon_t, \\ \sigma_t^2 = & \hat{\lambda}_0 + \hat{\lambda}_1 X_{t-1}^2 + \hat{\lambda}_2 \sigma_{t-1}^2, \end{cases} \quad (6.20)$$

Furthermore, we fit the $GARCH(1,1)$ model to our dataset, where the innovations ε_t are independent and identically Student- t distributed with $\hat{\lambda}_0 = 2.3 \times 10^{-05} (2.657 \times 10^{-05})$, $\hat{\lambda}_1 = 7.574 \times 10^{-01} (8.665 \times 10^{-01})$, $\hat{\lambda}_2 = 7.037 \times 10^{-01} (6.539 \times 10^{-02})$ and the parameter of the Student- t is $\hat{\nu} = 2.189 (2.453 \times 10^{-01})$,

where the value in parentheses is the standard deviation. Our aim is to estimate the Value-at-Risk of the return series at the 99.9% level, which corresponds to a extreme quantile with tail probability 0.1%, *i.e.*, $x(0.001)$.

Following the estimation procedure, we firstly estimate the second order parameter $\widehat{\rho}_{k_\rho} = -1.468$ which corresponds to the value of $k_\rho = 442$. Secondly, we use the four estimators which are $\widehat{\gamma}_{k_n}^{(H)}$, $\widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(dH)}$, $\widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(Ub1)}$ and $\widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(Ub2)}$ to estimate the tail index of the loss return series. We see that our $\widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(Ub2)}$ estimator seems to be more stable while the others increase with k_n . Finally, we use $\widehat{x}_{k_n}^{(W)}(p)$, $\widehat{x}_{k_n, \widehat{\rho}_{k_\rho}}^{(dH)}(p)$, $\widehat{x}_{k_n, \widehat{\rho}_{k_\rho}}^{(Ub1)}(p)$, and $\widehat{x}_{k_n, \widehat{\rho}_{k_\rho}}^{(Ub2)}(p)$ to estimate the VaR at 99% and 99.9% levels. It is clear that the $\widehat{x}_{k_n, \widehat{\rho}_{k_\rho}}^{(Ub2)}(p)$ quantile estimator is more stable than the others regarding the middle and right hand side Figure 6.4. That means it is the best estimator which can be used for high quantile estimation.

To illustrate that, we proceed by evaluating the optimal value of k_n that we denote by k_n^* . In the figure, this is the point around which a tail index estimator $\widehat{\gamma}_{k_n}^{(\bullet)}$ is stable. It is essential to decide a value of k_n which will be used to get the high quantile estimator. The selection of k_n^* is equivalent to the choice of the threshold in the EVT peaks-over-threshold method. An alternative method is the algorithm of (24), p.137, which allows to get an automatic choice of the number of top extremes k_n for a given tail index estimator $\widehat{\gamma}_{k_n}^{(\bullet)}$. According to these authors, an automatic choice of k_n^* is the k_n value which minimizes

$$\frac{1}{k_n} \sum_{j=1}^{k_n} j^\tau \left| \widehat{\gamma}_j^{(\bullet)} - \text{med} \left(\widehat{\gamma}_1^{(\bullet)}, \dots, \widehat{\gamma}_{k_n}^{(\bullet)} \right) \right|, \quad 1 \leq k_n \leq N,$$

where $0 \leq \tau < 1/2$ and $\text{med} \left(\widehat{\gamma}_1^{(\bullet)}, \dots, \widehat{\gamma}_{k_n}^{(\bullet)} \right)$ denotes the median of $\left(\widehat{\gamma}_1^{(\bullet)}, \dots, \widehat{\gamma}_{k_n}^{(\bullet)} \right)$ and N is the size of the negative log-return series.

By the way, choosing $\tau = 1/4$, we compute the optimal values k_n^* associated to the estimated values $\widehat{\gamma}_{k_n}^{(\bullet)} \in \left\{ \widehat{\gamma}_{k_n}^{(H)}, \widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(Ub1)}, \widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(Ub2)}, \widehat{\gamma}_{k_n, \widehat{\rho}_{k_\rho}}^{(dH)} \right\}$. In Table 6.1 below, we present the results of the estimated values $\widehat{\gamma}_{k_n^*}^{(\bullet)}$ with their associated extreme quantile estimators. Since we do not employ a parametric model for the time series, there is no explicit formula for calculating the asymptotic variance of the two estimators. Therefore, we opt to use a block bootstrapping method to construct the confidence interval for extreme quantiles estimators. The block bootstrapping follows the routine `boot` in the package `boot` in R software. By repeating such a bootstrapping procedure $K = 1000$ times, we obtain K bootstrapped estimates for each estimator. The sample standard deviation across the K estimates gives an estimate of the standard deviation of the underlying estimator for given $k_n \in \{1, \dots, N\}$. We construct the 95% confidence interval using the point estimate and the estimated standard deviation. This procedure is applied to all values of k_n of each estimator. The point estimates of the 99% and 99.9% Value at Risk as well as the lower and upper bounds of the confidence intervals are given in Table 6.1.

From Table 6.1, we remark that for moderate quantiles ($p = 0.01$), our $\widehat{x}_{k_n^*, \widehat{\rho}_{k_\rho}}^{(Ub2)}(p)$ estimator is not very competitive regarding the cover values. In contrast, for high quantiles ($p = 0.001$), our estimator is the best compared to that $\widehat{x}_{k_n^*, \widehat{\rho}_{k_\rho}}^{(dH)}(p)$. That illustrates well our conclusions drawn from the graphical analysis.

Tail index			Value at Risk						
$\hat{\gamma}_{k_n}^{(\bullet)}$	k_n^*	$\hat{\gamma}_{k_n^*}^{(\bullet)}$	$\hat{x}_{k_n^*}^{(\bullet)}(p)$	$p = 0.01$	95%-Conf. Int	Cover	$p = 0.001$	95%-Conf. Int	Cover
$\hat{\gamma}_{k_n}^{(H)}$	68	0.4665	$\hat{x}_{k_n^*}^{(W)}(p)$	0.0597	(0.0575, 0.1418)	0.0843	0.1752	(0.1602, 0.5949)	0.4347
$\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub1)}$	104	0.4005	$\hat{x}_{k_n^*, \hat{\rho}_{k_\rho}}^{(Ub1)}(p)$	0.0552	(0.0545, 0.1017)	0.0472	0.1393	(0.1322; 0.5487)	0.4165
$\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(Ub2)}$	126	0.3919	$\hat{x}_{k_n^*, \hat{\rho}_{k_\rho}}^{(Ub2)}(p)$	0.0511	(0.0129, 0.0594)	0.0465	0.1264	(0.0476, 0.1879)	0.1403
$\hat{\gamma}_{k_n, \hat{\rho}_{k_\rho}}^{(dH)}$	116	0.4006	$\hat{x}_{k_n^*, \hat{\rho}_{k_\rho}}^{(dH)}(p)$	0.0543	(0.0538, 0.0710)	0.0172	0.1369	(0.1302, 0.4352)	0.3050

TABLE 6.1. Estimated values of the Value at Risk with 95% confidence intervals at the optimal point of the SP500 negative log-return series. These values are computed with $\hat{\rho}_{k_\rho} = 1.468$ which corresponds to the value of $k_\rho = 442$.

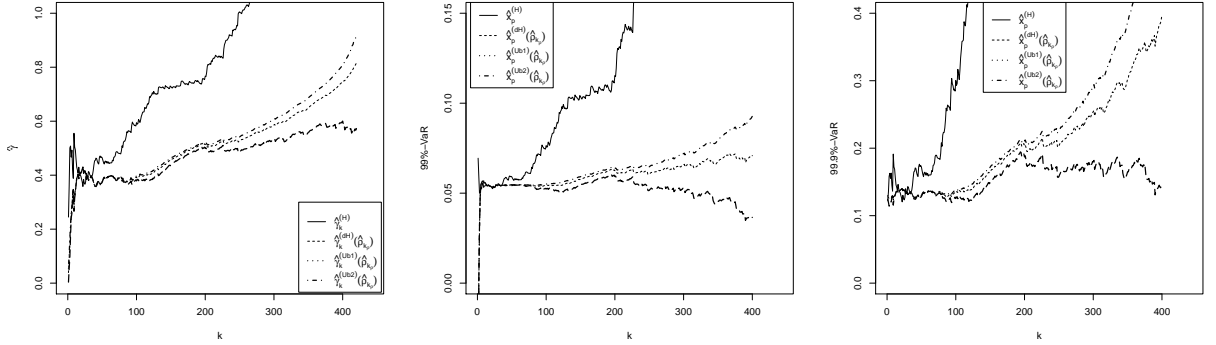


FIGURE 6.4. S&P500 index data: estimated values of γ (left panel), 99%-VaR (middle panel) and 99.9%-VaR (right panel) as function of k_n using the four (04) estimators.

7. PROOFS OF THE MAIN RESULTS

Proof of Proposition 2.1

It is similar to that of Corollary A.2 in (2) where we limit ourselves only to the second order conditions. Assume that the assumptions in Proposition 2.1 hold. With respect to the same Skorohod construction as in Proposition 1 in (16), it results that, for given $\alpha, \delta > 0$, there exist a function $\tilde{A} \sim A$, and a centered Gaussian process $(W(t))_{t \in [0,1]}$ with covariance function r , such that, as $n \rightarrow \infty$:

$$\sup_{t \in (0,1)} t^{1/2+\delta} \left| \sqrt{k_n} \left(\log \frac{Q_n(t)}{Q_n(1)} - \gamma(-\log t) \right) - \gamma(t^{-1}W(t) - W(1)) - \sqrt{k_n} \tilde{A}(n/k_n) \frac{t^{-\rho} - 1}{\rho} \right| \rightarrow 0 \text{ a.s.} \quad (7.21)$$

This means for all $t \in (0, 1]$

$$\left(\log \frac{Q_n(t)}{Q_n(1)} \right)^\alpha = \gamma^\alpha (-\log t)^\alpha \left[1 + \frac{(-\log t)^{-1}}{\sqrt{k_n}} (t^{-1}W(t) - W(1)) \right. \\ \left. + \gamma^{-1} (-\log t)^{-1} \tilde{A}(n/k_n) \frac{t^{-\rho} - 1}{\rho} + o(k_n^{-1/2}) t^{-1/2-\delta} \right]^\alpha.$$

And from the relation $(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + o(x^2)$, $o(x^2) \rightarrow 0$ when $x \rightarrow 0$, we obtain

$$\sup_{t \in (0,1]} t^{1/2+\delta} \left| \sqrt{k_n} \left(\log \frac{Q_n(t)}{Q_n(1)} \right)^\alpha - \gamma^\alpha (-\log t)^\alpha - \alpha \gamma^\alpha (-\log t)^{\alpha-1} (t^{-1}W(t) - W(1)) \right. \\ \left. - \sqrt{k_n} \tilde{A}(n/k_n) \alpha \gamma^{\alpha-1} (-\log t)^{\alpha-1} \frac{t^{-\rho} - 1}{\rho} \right| \rightarrow 0 \text{ a.s.}$$

Without losing generality, some terms tend to 0, as $n \rightarrow \infty$. In fact, we have $\sup_{t \in (0,1]} t^{1/2+\epsilon} t^{-1} |W(t)| = O(1)$ a.s. and $\tilde{A}(n/k_n) \rightarrow 0$ as $n \rightarrow \infty$.

Proof of Corollary 2.1

Obviously, we have

$$M_k^{(\alpha)} = \int_0^1 \left(\log \frac{Q_n(t)}{Q_n(1)} \right)^\alpha dt \quad (7.22)$$

By using $\delta < 1/2$ from the Proposition 2.1, we can take the integral of $\left(\log \frac{Q_n(t)}{Q_n(1)} \right)^\alpha$ on $(0, 1]$ and use the fact that $\int_0^1 (-\log t)^{a-1} t^{-b} dt = \frac{\Gamma(a)}{(1-b)^a}$ for $b < 1$ to obtain the result in the Corollary. The random term is obtained by taking

$$P^{(\alpha)} = \int_0^1 (-\log t)^{\alpha-1} (t^{-1}W(t) - W(1)) dt.$$

and is normally distributed with mean zero and covariance $Cov(P^{(\alpha)}, P^{(\beta)}) = c_{\alpha,\beta}$ defined as,

$$c_{\alpha,\beta} = \int_0^1 \int_0^1 (-\log s)^{\alpha-1} (-\log t)^{\beta-1} \times \left(\frac{r(s,t)}{st} - \frac{r(s,1)}{s} - \frac{r(1,t)}{t} + r(1,1) \right) ds dt,$$

with the covariance structure r defined as in the regularity condition (C_R) .

Proof of Theorem 2.1

With respect to the assumptions from Corollary 2.1 the following representation:

$$M_{k_n}^{(\alpha)} \stackrel{d}{=} \gamma^\alpha \Gamma(\alpha+1) + \frac{\alpha \gamma^\alpha P^{(\alpha)}}{\sqrt{k_n}} + \tilde{A}(n/k_n) \gamma^{\alpha-1} \frac{\Gamma(\alpha+1)}{\rho} \left(\frac{1}{(1-\rho)^\alpha} - 1 \right) + o_{\mathbb{P}}(k_n^{-1/2}) \quad (7.23)$$

It follows

$$\frac{M_{k_n}^{(\alpha)}}{\Gamma(\alpha+1)} \stackrel{d}{=} \gamma^\alpha \left(1 + \frac{\alpha P^{(\alpha)}}{\sqrt{k_n} \Gamma(\alpha+1)} + \tilde{A}(n/k_n) \frac{\gamma^{-1}}{\rho} \left(\frac{1}{(1-\rho)^\alpha} - 1 \right) + o_{\mathbb{P}}(k_n^{-1/2}) \right). \quad (7.24)$$

Therefore

$$\tilde{\gamma}_{k_n}^{(\alpha)} := \left(\frac{M_{k_n}^{(\alpha)}}{\Gamma(\alpha+1)} \right)^{\frac{1}{\alpha}} \stackrel{d}{=} \gamma \left(1 + \frac{\alpha P^{(\alpha)}}{\sqrt{k_n} \Gamma(\alpha+1)} + \tilde{A}(n/k_n) \frac{\gamma^{-1}}{\rho} \left(\frac{1}{(1-\rho)^\alpha} - 1 \right) + o_{\mathbb{P}}(k_n^{-1/2}) \right)^{\frac{1}{\alpha}}.$$

Using a Taylor expansion, we get as $n \rightarrow \infty$,

$$\begin{aligned}\widehat{\gamma}_{k_n}^{(\alpha)} &\stackrel{d}{=} \gamma \left(1 + \frac{1}{\alpha} \left[\frac{\alpha P^{(\alpha)}}{\sqrt{k_n} \Gamma(\alpha + 1)} + \tilde{A}(n/k_n) \frac{\gamma^{-1}}{\rho} \left(\frac{1}{(1-\rho)^\alpha} - 1 \right) \right] + o_{\mathbb{P}}(k_n^{-/2}) \right) \\ &\stackrel{d}{=} \gamma + \frac{\gamma P^{(\alpha)}}{\sqrt{k_n} \Gamma(\alpha + 1)} + \frac{\tilde{A}(n/k_n)}{\alpha \rho} \left(\frac{1}{(1-\rho)^\alpha} - 1 \right) + o_{\mathbb{P}}(k_n^{-/2}).\end{aligned}$$

The expression (2.5) of $\widehat{\gamma}_{k_n}^{(\alpha)}$ can be written as:

$$\widehat{\gamma}_{k_n}^{(\alpha)} = \frac{M_{k_n}^{(\alpha)}}{\Gamma(\alpha + 1)} \left(M_{k_n}^{(1)} \right)^{1-\alpha}, \quad \alpha > 1.$$

From (7.23), it follows:

$$\begin{aligned}M_{k_n}^{(1)} &\stackrel{d}{=} \gamma + \frac{\gamma P^{(1)}}{\sqrt{k_n}} + \frac{\tilde{A}(n/k_n)}{1-\rho} + o_{\mathbb{P}}(k_n^{-/2}) \\ &\stackrel{d}{=} \gamma \left[1 + \frac{P^{(1)}}{\sqrt{k_n}} + \tilde{A}(n/k_n) \frac{\gamma^{-1}}{1-\rho} + o_{\mathbb{P}}(k_n^{-/2}) \right]. \\ \left(M_{k_n}^{(1)} \right)^{1-\alpha} &\stackrel{d}{=} \gamma^{1-\alpha} \left(1 + \frac{P^{(1)}}{\sqrt{k_n}} + \tilde{A}(n/k_n) \frac{\gamma^{-1}}{1-\rho} + o_{\mathbb{P}}(k_n^{-/2}) \right)^{1-\alpha} \\ &\stackrel{d}{=} \gamma^{1-\alpha} \left(1 + \frac{(1-\alpha)P^{(1)}}{\sqrt{k_n}} + \tilde{A}(n/k_n) \gamma^{-1} \frac{1-\alpha}{1-\rho} + o_{\mathbb{P}}(k_n^{-/2}) \right).\end{aligned}$$

We recall that

$$\frac{M_{k_n}^{(\alpha)}}{\Gamma(\alpha + 1)} \stackrel{d}{=} \gamma^\alpha \left(1 + \frac{\alpha P^{(\alpha)}}{\sqrt{k_n} \Gamma(\alpha + 1)} + \tilde{A}(n/k_n) \frac{\gamma^{-1}}{\rho} \left(\frac{1}{(1-\rho)^\alpha} - 1 \right) + o_{\mathbb{P}}(k_n^{-/2}) \right).$$

Hence

$$\widehat{\gamma}_{k_n}^{(\alpha)} \stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k_n}} \widehat{P}^{(\alpha)} + \tilde{A}(n/k_n) \left\{ \frac{1 - (1-\rho)^\alpha}{\rho(1-\rho)^\alpha} - \frac{\alpha - 1}{1-\rho} \right\} + o_{\mathbb{P}}(k_n^{1/2}).$$

Where $\widehat{P}^{(\alpha)} = \frac{\alpha P^{(\alpha)}}{\Gamma(\alpha+1)} - (\alpha - 1)P^{(1)}$, which ends the proof of the Theorem 2.1.

Proof of Corollary 2.2

The results in Corollary 2.2 quite direct by computing the variance of the centered Gaussian term $P^{(\alpha)}$ given in Corollary 2.1.

Proof of Theorem 2.2

Note That:

$$\widehat{\gamma}_{k_n}^{(H)} = \widehat{\gamma}_{k_n}^{(1)} = \widehat{\gamma}_{k_n}^{(1)} = M_{k_n}^{(1)}, \quad \widehat{\gamma}_{k_n}^{(2)} = \frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} \quad \text{and} \quad \widetilde{\gamma}_{k_n}^{(2)} = \sqrt{\frac{M_{k_n}^{(2)}}{2}}. \quad (7.25)$$

We recall that:

$$\begin{aligned}\widehat{\gamma}_{k_n, \widehat{\rho}}^{(Ub1)} &:= \frac{1}{\widehat{\rho}} \left(\sqrt{2M_{k_n}^{(2)}} - (2 - \widehat{\rho}) \frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} \right), \\ \widehat{\gamma}_{k_n, \widehat{\rho}}^{(Ub2)} &:= \frac{1}{\widehat{\rho}} \left((2 - \widehat{\rho}) M_{k_n}^{(1)} - (1 - \widehat{\rho}) \sqrt{2M_{k_n}^{(2)}} \right)\end{aligned}$$

and

$$\widehat{\gamma}_{k_n, \widehat{\rho}}^{(dH)} := \frac{1}{\widehat{\rho}} \left(M_{k_n}^{(1)} - (1 - \widehat{\rho}) \frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} \right),$$

where $\widehat{\rho}$ is either a canonical negative value $\widehat{\rho} := \rho = \rho_0$ or a consistent estimator $\widehat{\rho} := \widehat{\rho}_{k_\rho}$ of ρ , with $k_\rho := k_{\rho, n}$ an intermediate sequence of integers greater than k_n , satisfying $k_\rho \rightarrow \infty$ and $k_\rho/n \rightarrow 0$, as $n \rightarrow \infty$.

From Theorem 2.1, we have:

$$M_{k_n}^{(1)} \stackrel{d}{=} \gamma + \frac{\gamma P^{(1)}}{k_n^{1/2}} + \frac{\widetilde{A}(n/k_n)}{1 - \rho} + o_{\mathbb{P}}(k_n^{-1/2}), \quad (7.26)$$

$$\sqrt{2M_{k_n}^{(2)}} \stackrel{d}{=} 2\sqrt{\frac{M_{k_n}^{(2)}}{2}} \stackrel{d}{=} 2\gamma + \frac{\gamma P^{(2)}}{k_n^{1/2}} + \widetilde{A}(n/k_n) \frac{2 - \rho}{(1 - \rho)^2} + o_{\mathbb{P}}(k_n^{-1/2}), \quad (7.27)$$

and

$$\frac{M_{k_n}^{(2)}}{2M_{k_n}^{(1)}} \stackrel{d}{=} \gamma + \frac{\gamma P^{(2)}}{k_n^{1/2}} - \frac{\gamma P^{(1)}}{k_n^{1/2}} + \widetilde{A}(n/k_n) \frac{1}{(1 - \rho)^2} + o_{\mathbb{P}}(k_n^{-1/2}). \quad (7.28)$$

Using the Jackknife methodology, this leads to:

$$\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} \stackrel{d}{=} \gamma + \frac{\gamma(2 - \widehat{\rho})P^{(1)} - \gamma(1 - \widehat{\rho})P^{(2)}}{\widehat{\rho}\sqrt{k_n}} + \widetilde{A}(n/k) \frac{\widehat{\rho} - \rho}{\widehat{\rho}(1 - \rho)^2} + o_{\mathbb{P}}(k_n^{-1/2}).$$

For a given canonical negative value $\widehat{\rho} := \rho = \rho_0$ or a consistent estimator $\widehat{\rho} := \widehat{\rho}_{k_\rho}$ of ρ , we get as $n \rightarrow \infty$

$$k_n^{1/2} \left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma \right) \xrightarrow{d} \frac{\gamma}{\rho} \left((2 - \rho)P^{(1)} - (1 - \rho)P^{(2)} \right).$$

Computing the the variance term of $\frac{\gamma}{\rho} \left((2 - \rho)P^{(1)} - (1 - \rho)P^{(2)} \right)$ with respect to the covariance structure, the Theorem 2.2 holds.

Proof of Corollary 2.2

According to (2), under (C_{SO}) and (C_R) assumptions, if $k_\rho := k_{\rho, n}$ satisfies $k_\rho \rightarrow \infty$, $k_\rho/n \rightarrow 0$ and $\sqrt{k_\rho}A(n/k_\rho) \rightarrow \infty$, as $n \rightarrow \infty$, then $\widehat{\rho}_{k_\rho}^{(*)} \xrightarrow{\mathbb{P}} \rho$. The Corollary 2.3 follows by applying the Theorem 2.2.

Proof of Theorem 2.3

For simplify, let's denote by $d_n = \frac{k_n}{n\rho}$ and

$$T_n = \frac{\left(M_{k_n}^{(2)} - 2(M_{k_n}^{(1)})^2 \right) (1 - \widehat{\rho})^2}{2M_{k_n}^{(1)} \widehat{\rho}^2} \left(1 - d_n^{\widehat{\rho}} \right).$$

We consider the expression:

$$E_n = \frac{\sqrt{k_n}}{\log d_n} \left(\frac{\widehat{x}_{k_n, \widehat{\rho}}^{(\bullet)}(p)}{x(p)} - 1 \right) \quad \text{with} \quad \widehat{x}_{k_n, \widehat{\rho}}^{(\bullet)}(p) = X_{n-k_n, n} d_n^{\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)}} (1 - T_n)$$

. Let $x(p) = U(1/p)$. It follows

$$\begin{aligned}
E_n &= \frac{\sqrt{k_n}}{\log d_n} \left(\frac{X_{n-k_n, n} \widehat{d}_{k_n, \hat{\rho}}^{(\bullet)}}{x(p)} (1 - T_n) - 1 \right) \\
&= \frac{\sqrt{k_n}}{\log d_n} \left[\left(\frac{X_{n-k_n, n} \widehat{d}_{k_n, \hat{\rho}}^{(\bullet)}}{x(p)} - 1 + 1 \right) (1 - T_n) - 1 \right] \\
&= \frac{\sqrt{k_n}}{\log d_n} \left[\left(\frac{X_{n-k_n, n} \widehat{d}_{k_n, \hat{\rho}}^{(\bullet)}}{x(p)} - 1 \right) (1 - T_n) + (1 - T_n) - 1 \right] \\
&= \frac{\sqrt{k_n}}{\log d_n} \left(\frac{X_{n-k_n, n} \widehat{d}_{k_n, \hat{\rho}}^{(\bullet)}}{x(p)} - 1 \right) (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} T_n \\
&= \frac{\sqrt{k_n}}{\log d_n} \frac{X_{n-k_n, n} \widehat{d}_{k_n, \hat{\rho}}^{(\bullet)}}{x(p)} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} T_n \\
&= \frac{\sqrt{k_n}}{\log d_n} \frac{X_{n-k_n, n} \widehat{d}_{k_n, \hat{\rho}}^{(\bullet) - \gamma + \gamma} U(n/k_n)}{x(p) U(n/k_n)} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} T_n \\
&= \frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} \frac{\sqrt{k_n}}{\log d_n} \frac{X_{n-k_n, n} \widehat{d}_{k_n, \hat{\rho}}^{(\bullet) - \gamma}}{U(n/k_n)} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} T_n \\
&= \frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} \left[\frac{\sqrt{k_n}}{\log d_n} \left(\frac{X_{n-k_n, n}}{U(n/k_n)} - 1 \right) \widehat{d}_{k_n, \hat{\rho}}^{(\bullet) - \gamma} + \frac{\sqrt{k_n}}{\log d_n} \widehat{d}_{k_n, \hat{\rho}}^{(\bullet) - \gamma} \right] (1 - T_n) \\
&\quad - \frac{\sqrt{k_n}}{\log d_n} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} T_n \\
&= \frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} \left[\frac{\sqrt{k_n}}{\log d_n} \left(\frac{X_{n-k_n, n}}{U(n/k_n)} - 1 \right) \widehat{d}_{k_n, \hat{\rho}}^{(\bullet) - \gamma} + \frac{\sqrt{k_n}}{\log d_n} \left(\widehat{d}_{k_n, \hat{\rho}}^{(\bullet) - \gamma} - 1 \right) \right] (1 - T_n) \\
&\quad + \frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} \frac{\sqrt{k_n}}{\log d_n} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n}.
\end{aligned}$$

We have as $n \rightarrow \infty$, $\frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} \rightarrow 1$, because $\frac{U(1/p_n) d_n^{-\gamma}}{U(n/k_n)} \rightarrow 1$, by the regularly varying condition of the tail quantile U .

From Proposition 1. in (16), we have as $n \rightarrow \infty$,

$$\sqrt{k} \left(\frac{X_{n-k, n}}{U(n/k)} - 1 \right) \xrightarrow{d} \gamma W(1). \quad (7.29)$$

Next, we also

$$\widehat{d}_{k_n, \hat{\rho}}^{(\bullet) - \gamma} = \exp \left[\left(\widehat{\gamma}_{k_n, \hat{\rho}}^{(\bullet)} - \gamma \right) \log d_n \right].$$

From Theorem 2.2, we get for all large values of n ;

$$\widehat{\gamma}_{k_n, \hat{\rho}}^{(\bullet)} - \gamma \stackrel{d}{=} k_n^{-1/2} \Gamma + o_{\mathbb{P}}(k^{-1/2})$$

as $n \rightarrow \infty$ where Γ is a centred Gaussian distribution with variance by assumption $\sigma^2(\gamma, \rho)$. Since $\log d_n/k_n^{1/2} \rightarrow 0$, as $n \rightarrow \infty$, this implies that for all n large enough,

$$\left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma\right) \log d_n = o_{\mathbb{P}}(1). \quad (7.30)$$

Further, by using the inequality, $\left|\frac{\exp(x)-1}{x} - 1\right| \leq \exp(|x|) - 1$, in the neighborhood of zero, we get

$$d_n^{\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma} \stackrel{d}{=} 1 + \left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma\right) \log d_n + O_{\mathbb{P}}(R_n), \quad (7.31)$$

as $n \rightarrow \infty$, with

$$R_n = \left| \left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma\right) \log d_n \right| \times \left[\exp \left\{ \left| \left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma\right) \log d_n \right| \right\} - 1 \right].$$

Since $np_n/k_n \rightarrow 0$, as $n \rightarrow \infty$, we have $d_n = k_n/(np_n) \rightarrow \infty$. Therefore, from (7.29), (7.30) and (7.31), we get as $n \rightarrow \infty$:

$$\frac{\sqrt{k_n}}{\log d_n} \left(\frac{X_{n-k_n, n}}{U(n/k_n)} - 1 \right) d_n^{\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma} = o_{\mathbb{P}}(1).$$

And in the other hand, we get as $n \rightarrow \infty$:

$$\frac{\sqrt{k_n}}{\log d_n} \left(d_n^{\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma} - 1 \right) \stackrel{d}{=} \sqrt{k_n} \left(\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma \right) + o_{\mathbb{P}}(1).$$

This leads to

$$\frac{\sqrt{k_n}}{\log d_n} \left(d_n^{\widehat{\gamma}_{k_n, \widehat{\rho}}^{(\bullet)} - \gamma} - 1 \right) \xrightarrow{d} \Gamma,$$

as $n \rightarrow \infty$. Next, from (7.23), $M_{k_n}^{(1)} \xrightarrow{\mathbb{P}} \gamma$ and $M_{k_n}^{(2)} \xrightarrow{\mathbb{P}} 2\gamma^2$, as $n \rightarrow \infty$. For a given canonical negative value $\widehat{\rho} := \rho = \rho_0$ or a consistent estimator $\widehat{\rho} := \widehat{\rho}_{k_\rho}$ of ρ , we get $T_n \xrightarrow{\mathbb{P}} 0$, as $n \rightarrow \infty$.

For the last term, we have:

$$\begin{aligned} F_n &= \frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} \frac{\sqrt{k_n}}{\log d_n} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} \\ &= \frac{\sqrt{k_n}}{\log d_n} \frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} \\ &= \frac{\sqrt{k_n}}{\log d_n} \left[\frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} - 1 \right] (1 - T_n) + \frac{\sqrt{k_n}}{\log d_n} (1 - T_n) - \frac{\sqrt{k_n}}{\log d_n} \\ &= \frac{\sqrt{k_n}}{\log d_n} \left[\frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} - 1 \right] (1 - T_n) + \frac{\sqrt{k_n}}{\log d_n} T_n \\ &= F_{n,1} + F_{n,2}. \end{aligned}$$

By expanding $F_{n,1}$ we get:

$$\begin{aligned} F_{n,1} &= \frac{\sqrt{k}}{\log d_n} \frac{d_n^\gamma U(n/k_n)}{U(1/p_n)} \left[1 - \frac{d_n^{-\gamma} U(1/p_n)}{U(n/k)} \right] (1 - T_n) \\ &= \frac{1}{\log d_n} \frac{-d_n^\gamma U(n/k)}{U(1/p_n)} \frac{\sqrt{k} A(n/k)}{\left[\frac{d_n^{-\gamma} U(1/p_n)}{U(n/k)} - 1 \right]} (1 - T_n). \end{aligned}$$

Since $\frac{-d_n^\gamma U(n/k)}{U(1/p_n)} \rightarrow 1$, $\sqrt{k}A(n/k) \rightarrow \lambda$, $\log d_n \rightarrow \infty$, $\frac{d_n^{-\gamma} U(1/p_n) - 1}{A(n/k)} \rightarrow -\frac{1}{\rho}$, and $T_n \xrightarrow{\mathbb{P}} 0$, we get $F_{n,1} \xrightarrow{\mathbb{P}} 0$, as $n \rightarrow \infty$.

Now, for $F_{n,2}$, we obtain:

$$F_{n,2} = \frac{\sqrt{k_n}}{\log d_n} \left(T_n - \frac{\tilde{A}(n/k_n)}{\rho} (1 - d_n^{\hat{\rho}}) \right) + \frac{\sqrt{k_n}}{\log d_n} \frac{\tilde{A}(n/k_n)}{\rho} (1 - d_n^{\hat{\rho}}).$$

For a given canonical negative value $\hat{\rho} := \rho = \rho_0$ or a consistent estimator $\hat{\rho} := \hat{\rho}_{k_\rho}$ of ρ , we get $d_n^{\hat{\rho}} \rightarrow 0$, as $n \rightarrow \infty$. Since $\sqrt{k}A(n/k) \rightarrow \lambda$, $\log d_n \rightarrow \infty$, we have, as $n \rightarrow \infty$:

$$\frac{\sqrt{k_n}}{\log d_n} \frac{\tilde{A}(n/k_n)}{\rho} (1 - d_n^{\hat{\rho}}) \rightarrow 0.$$

On the other hand

$$\frac{\sqrt{k_n}}{\log d_n} \left(T_n - \frac{\tilde{A}(n/k_n)}{\rho} (1 - d_n^{\hat{\rho}}) \right) = \frac{\sqrt{k_n}}{\log d_n} \left(\frac{(M_{k_n}^{(2)} - 2(M_{k_n}^{(1)})^2) (1 - \hat{\rho})^2}{2M_{k_n}^{(1)} \hat{\rho}^2} - \frac{\tilde{A}(n/k_n)}{\rho} \right) (1 - d_n^{\hat{\rho}}).$$

From (7.26) and (7.28), we have:

$$\sqrt{k_n} \left(\frac{M_{k_n}^{(2)} - 2(M_{k_n}^{(1)})^2}{2\hat{\rho}M_{k_n}^{(1)}} - \frac{\tilde{A}(n/k_n)}{\rho} \right) \xrightarrow{d} \frac{(1 - \rho)^2}{\rho^2} (P^{(2)} - 2P^{(1)}).$$

This implies that

$$\frac{\sqrt{k_n}}{\log d_n} \left(T_n - \frac{\tilde{A}(n/k_n)}{\rho} \right) \xrightarrow{\mathbb{P}} 0.$$

Finally, we get, as $n \rightarrow \infty$,

$$\frac{\sqrt{k_n}}{\log d_n} \left(\frac{\hat{x}_{k_n, \hat{\rho}}^{(\bullet)}(p)}{x(p)} - 1 \right) \xrightarrow{d} \Gamma.$$

This ends the proof of Theorem 2.3.

REFERENCES

- BEIRLANT J., DIERCKX G., GOEGBEUR Y., AND MATTHYS G. (1999). *Tail index estimation and an exponential regression model*. *Extremes* 2, 177 – 200.
- DE HAAN, L., MERCADIER, C. AND ZHOU, C. (2016). *Adapting extreme value statistics to financial time series: dealing with bias and serial dependence*. *Finance and Stochastics* 20, 321–354.
- DEKKERS, A. L. M., EINMAHL, J. H. J., DE HAAN, L. (1989). *A moment estimator for the index of an extreme-value distribution*. *Ann. Statist.* 17, 1833 – 1855.
- DEME, E. H., GARDES, L., AND GIRARD, S. (2013). *On the estimation of the second order parameter for heavy-tailed distributions*. *Statistical Journal* 3(11), 277 – 299.
- DREES H. (2000). *Weighted approximations of tail processes for β -mixing random variables*. *Annals of Applied Probability* 10, 1274–1301.

- DREES H. (2003). *Extreme quantile estimation for dependent data, with applications to finance*. Bernoulli 9, 617–657.
- DREES H., KAUFMANN E.(1998). *Selecting the optimal sample fraction in univariate extreme value estimation*. In: Stochastic Processes and their Applications, vol. 75, 149–172.
- FEUERVERGER A., HALL P.(1999). *Estimating a tail exponent by modelling departure from a Pareto distribution*. Annals of Statistics 27, 760–781.
- FRAGA ALVES, M., GOMES, M., AND DE HAAN, L.(2003). *A new class of semi-parametric estimators of the second order parameter*. Portugaliae Mathematica 60, 193–214.
- GOEGEBEUR, Y., BEIRLANT, J., AND DE WET, T.(2010). *Kernel estimators for the second order parameter in extreme value statistics*. Journal of Statistical Planning and Inference 140, 2632–2652.
- GOMES M.I., DE HAAN L., RODRIGUES L.(2008). *Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses*. J. R. Stat. Soc., Ser. B, Stat. Methodol. 70, 31–52.
- GOMES M.I., MARTINS M.J., AND NEVES M.(2000). *Alternatives to a semi-parametric estimator of parameters of rare events- The Jackknife methodology*, Extremes(3), 207 – 229.
- GOMES M.I., MARTINS M.J.(2001). *Generalizations of the Hill estimator — asymptotic versus finite sample behaviour*, J. Statistical Planning and Inference, 93, 161–180.
- GOMES M.I., MARTINS M.J.(2002). *”Asymptotically Unbiased” Estimators of the Tail Index Based on the Estimation of the Second Order Parameter*. Extremes, 5, 1, 5 – 31.
- GOMES M.I., MARTINS M.J.(2003). *Bias Reduction and Explicit Efficient Estimation of the Tail Index*, Notas e Comunicações C.E.A.U.L. 6/2002.
- GUILLOU A., C.-DE MOULIN V.(2018). *Extreme quantile estimation for β -mixing time series and applications*. Insurance : Mathematics and Economics, 59 – 74.
- HILL, B.M.(1975). *A simple general approach to inference about the tail of a distribution*. Ann. Statist. 3 : 1163 – 1174.
- HSING T. (1991). *On tail index estimation using dependent data*. Annals of Statistics 19, 1547–1569.
- JORDANOVA, P., FABIAAN, Z., HERMANN, P., STRELEC, L., RIVERA, A., GIRARD, S., TORRES, S., STEHLIK, M. (2016). *Weak properties and robustness of t -Hill estimators*. Extremes.
- KESTEN H.(1973). *Random difference equations and renewal theory for products of random matrices*. Acta Math. 131, 207–248.
- MATTHYS, G., DELAFOSSE, E., GUILLOU, A., AND BEIRLANT, J. (2004). *Estimating catastrophic quantile levels for heavy-tailed distributions*. Insurance: Mathematics and Economics 34, 517–537.
- MCNEIL, A., FREY, R. & EMBRECHTS, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools, Revised Edition*. Princeton University Press, Princeton, New Jersey.
- PENG L.(1998). *Asymptotically unbiased estimator for the extreme-value index*, Statistics and Probability Letters 38(2), 107 – 115.
- REISS, R.D., THOMAS, M. (2007). *Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. 3rd ed. Birkhäuser.

- STARICA, C. (1999). *On the tail empirical process of solutions of stochastic difference equations*. Technical report, Chalmers University and University of Pennsylvania.
- TARULLO, D.(2008). *Banking on basel. The future of international financial regulation*. Technical report, Peterson Institute for International Economics, Washington, DC.
- WEISSMAN, I. (1978). *Estimation of parameters and large quantiles based on the k largest observations*. J. Am. Statist. Assoc. 73 : 812 – 815.
- WORMS, J. AND WORMS, R. (2011). *Estimation of second order parameters using probability weighted moments*. Probability and Statistics 16, 97 – 113.