Adaptive PORT-MVRB Estimation: an Empirical Comparison of two Heuristic Algorithms*

M. Ivette Gomes

Universidade de Lisboa, DEIO, CEAUL and FCUL

Lígia Henriques-Rodrigues

Instituto Politécnico de Tomar, and CEAUL

Abstract. In this article, we deal with an empirical comparison of two data-driven heuristic procedures of estimation of a positive *extreme value index* (EVI), working thus with heavy right tails. The semi-parametric EVI-estimators under consideration, the so-called PORT-MVRB EVI-estimators, are location and scale-invariant estimators, based on the *peaks over random threshold* (PORT) methodology applied to second-order *minimum-variance reduced-bias* (MVRB) EVI-estimators. Due to the stability on k of the MVRB EVI-estimates, as well as of the PORT-MVRB-estimates, for adequate values of q, we consider the use of two heuristic algorithms, both based on bias stability. Trivial adaptations of these algorithms make them work for a similar estimation of other parameters of *extreme events*, like the *Value-at-Risk* at the level p, the size of the loss occurred with a small probability p, and the *probability of exceedance* of a high level x, among others. Applications to simulated data sets and to real data sets in the field of finance are provided.

*Research partially supported by FCT / OE and PTDC / FEDER.

AMS 2000 subject classification. Primary 62G32, 62E20; Secondary 65C05.

Keywords and phrases. *Statistics of extremes; extreme value index; semi-parametric estimation; adaptive choices; heuristics; bias reduction; location/scale invariant estimation.*

1 Introduction and scope of the article

For heavy right tails, i.e., for models F in the domain of attraction for maxima of an *extreme value* (EV) distribution function (d.f.),

$$EV_\gamma(x) = \exp\big(-(1+\gamma x)^{-1/\gamma}\big),\ x > -1/\gamma, \tag{1.1}$$

with $\gamma(>0)$ the so-called *extreme value index* (EVI), we first refer the classical EVI-estimators, the Hill estimators, derived and studied in Hill (1975). Given a random sample $\underline{\mathbf{X}}_n = (X_1, \ldots, X_n)$ and the associated sample of ascending order statistics (o.s.'s), $(X_{1:n} \le \cdots \le X_{n:n})$, Hill estimators are defined for $k = 1, 2, \ldots, n-1$, and are given by

$$H(k) \equiv H(k; \underline{\mathbf{X}}_n) := \frac{1}{k}\sum_{i=1}^{k}\{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}, \tag{1.2}$$

being consistent for the estimation of γ provided that k is *intermediate*, i.e., provided that

$$k = k_n \to \infty \quad \text{and} \quad k/n \to 0,\ \text{as } n \to \infty. \tag{1.3}$$

Due to the high bias of the Hill estimators, in (1.2), for moderate up to large k, several authors have been dealing with bias reduction in the field of extremes. We refer the pioneering articles by Beirlant *et al.* (1999), Feuerveger and Hall (1999) and Gomes *et al.* (2000), as well as the most recent second-order *minimum-variance reduced-bias* (MVRB) EVI-estimators in Caeiro *et al.* (2005) and Gomes *et al.* (2007, 2008b). The simplest class of second-order MVRB EVI-estimators is the one in Caeiro *et al.* (2005). This class, denoted $\overline{H} \equiv \overline{H}(k)$, depends upon the estimation of the second-order parameters (β, ρ), $\beta \neq 0$, $\rho < 0$, in the function $A(t) = \gamma\beta t^\rho$, which measures, in an adequate sense, the rate of convergence of the sequence of maximum values, $X_{n:n}$, linearly normalized, towards a non-degenerate random variable (r.v), with a d.f. necessarily of the type of the EV d.f. in (1.1). With the notation $F^{\leftarrow}(x) := \inf\{y : F(y) \geq x\}$ for the generalized inverse function of F, the function $A(\cdot)$ appears then in the *reciprocal quantile function* $U(t) := F^{\leftarrow}(1 - 1/t)$, which can be written as

$$U(t) = Ct^\gamma \left(1 + A(t)/\rho + o(t^\rho)\right), \quad A(t) = \gamma\beta t^\rho. \tag{1.4}$$

The functional form of those EVI-estimators is

$$\overline{H}(k) \equiv \overline{H}(k; \underline{\mathbf{X}}_n) \equiv \overline{H}_{\hat\beta,\hat\rho}(k) := H(k)\big(1 - \hat\beta(n/k)^{\hat\rho}/(1-\hat\rho)\big), \tag{1.5}$$

with $H(k)$ the Hill estimator in (1.2), and where $(\hat\beta, \hat\rho)$ needs to be any adequate consistent estimator of (β, ρ).

The estimators in (1.2) and (1.5) are scale-invariant, but not location-invariant, as often desired, and this contrarily to the PORT-Hill estimators, introduced in Araújo Santos *et al.* (2006) and further studied in Gomes *et al.* (2008a), with PORT standing for *peaks over random thresholds*. The class of PORT-Hill estimators is based on a *sample of excesses* over a random threshold $X_{n_q:n}$, $n_q := [nq] + 1$, where $[x]$ denotes, as usual, the integer part of x, i.e., it is based on

$$\underline{\mathbf{X}}_n^{(q)} := \big(X_{n:n} - X_{n_q:n}, \ldots, X_{n_q+1:n} - X_{n_q:n}\big). \tag{1.6}$$

We can generally have $0 < q < 1$, for d.f.'s with finite or infinite left endpoint $x_{_F} := \inf\{x : F(x) > 0\}$ (*the random threshold is an empirical quantile*), and $q = 0$, for d.f.'s with finite left endpoint $x_{_F}$ (*the random threshold is the minimum*). In this article, we shall work with the EVI-estimators, already considered in Gomes *et al.* (2010a), the so-called PORT-MVRB estimators. They have the same functional form of the MVRB estimators in (1.5), but with the original sample $\underline{\mathbf{X}}_n$ replaced everywhere by the sample of excesses $\underline{\mathbf{X}}_n^{(q)}$, in (1.6). Consequently, with $(\hat\beta_q, \hat\rho_q)$ any adequate estimator of (β, ρ), based on the sample $\underline{\mathbf{X}}_n^{(q)}$, such estimators are given by the functional equation,

$$\overline{H}^{(q)}(k) := H(k; \underline{\mathbf{X}}_n^{(q)})\Big(1 - \hat\beta_q\ (n/k)^{\hat\rho_q}/(1 - \hat\rho_q)\Big), \tag{1.7}$$

with $H(k; \underline{\mathbf{X}}_n)$ and $\underline{\mathbf{X}}_n^{(q)}$ given in (1.2) and (1.6), respectively. These estimators are now invariant for both changes of location and scale, and depend on the *tuning parameter* q, which only influences their asymptotic bias, making them highly flexible, and able to compare favourably with the MVRB estimators in (1.5), for a large variety of underlying models in the domain of attraction for maxima of the EV d.f., in (1.1). See Gomes *et al.* (2010a).

Section 2 of this article is dedicated to data-driven choices of the *tuning parameters* under play, after a brief review, in Section 2.1, of the estimation of the second-order parameters (β, ρ), in (1.4). In Section 2.2, due to the reasonably high stability on k of the MVRB estimates $\overline{H}$, in (1.5), and $\overline{H}^{(q)}$, in (1.7), for adequate values of q, we propose a first algorithm for the choice of k, whenever we use $\overline{H}$, in (1.5), as well as for the choice of k and q, whenever we use $\overline{H}^{(q)}$, in (1.7), as estimators of the EVI. Such algorithm is based on the bias pattern of the estimators themselves as a function of k, and it is similar to the ones used in Figueiredo *et al.* (2010) for a data-driven estimation of the Value-at-Risk, and in Gomes *et al.* (2010a), for a data-driven choice of the EVI. We have also decided for a second algorithm, described in Section 2.3, easier to implement computationally than the one in Section 2.2. Such an algorithm is also based on bias patterns of the estimates, but after a simple smoothing operation, which cuts the small "ups" and "downs" in the stability region of the estimates, as function of k, simplifying the identification of "runs". In Section 2.3.1, we illustrate the behaviour of these modified estimators through a Monte-Carlo simulation related with a Student underlying parent. Finally, in Section 3, and prior to some final comments, in Section 3.6, we provide applications of the adaptive methodologies to data in the fields of finance, as well as to simulated samples from a Student underlying parent. The illustration of the performance of the algorithms is done here through the EVI-estimation, but trivial adaptations make them work for the associated estimation of other parameters of *extreme events*, like the *Value-at-Risk* at the level p, the size of the loss occurred with a small probability p, and the *probability of exceedance* of a high level x, among others.

2 Second-order parameters estimation and data-driven choices of tuning parameters

Similarly to Figueiredo et al. (2010) and Gomes *et al.* (2010a), we shall use the notation $X_{0:n} = 0$, so that with $\overline{H}$ and $\overline{H}^{(q)}$, given in (1.5) and (1.7), respectively, we can consider that $\overline{H} = \overline{H}^{(q)}$ for $q = -1/n$ ($n_q = 0$). Our interest lies then on the estimation of γ through $\overline{H}^{(q)}$, in (1.7), also including $\overline{H}$, in (1.5). We first review, in Section 2.1, the kind of estimation of second-order parameters used in this article. Next, in Section 2.2, a heuristic method of the type of the ones proposed in Figueiredo et al. (2010) and Gomes *et al.* (2010a), is written algorithmically. In order to simplify the computational counting of "*largest runs*", we next propose, in Section 2.3, a second adaptive heuristic estimation of γ, where we slightly modify the pattern of the PORT-MVRB EVI-estimates, as function of k. On the basis of a small-scale Monte-Carlo simulation, the main properties of such a modified EVI-estimator are provided in Section 2.3.1.

2.1 Estimation of second-order parameters

Up to now, the most commonly used estimators of β and ρ, also considered in this article, are the ones introduced in Gomes and Martins (2002) and Fraga Alves *et al.* (2003), respectively, briefly introduced in the sequel. Given the sample $\underline{\mathbf{X}}_n$, the

ρ-estimators in Fraga Alves *et al.* (2003) are dependent on the statistics

$$T_\tau(k;\underline{\mathbf{X}}_n) := \frac{\left(M_n^{(1)}(k;\underline{\mathbf{X}}_n)\right)^\tau - \left(M_n^{(2)}(k;\underline{\mathbf{X}}_n)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k;\underline{\mathbf{X}}_n)/2\right)^{\tau/2} - \left(M_n^{(3)}(k;\underline{\mathbf{X}}_n)/6\right)^{\tau/3}}, \tag{2.1}$$

defined for any *tuning parameter* $\tau \in \mathbb{R}$, provided that we consider the usual notation $a^{b\tau} = b\ln a$ if $\tau = 0$, and where

$$M_n^{(j)}(k) := \frac{1}{k}\sum_{i=1}^{k}\left\{\ln X_{n-i+1:n} - \ln X_{n-k:n}\right\}^j, \quad j = 1, 2, 3.$$

Under mild restrictions on k, the statistics in (2.1) converge towards $3(1-\rho)/(3-\rho)$, independently of the *tuning* parameter τ, and we consequently have the ρ-estimators,

$$\hat\rho_\tau(k) \equiv \hat\rho_\tau(k;\underline{\mathbf{X}}_n) := \min\left(0,\ \frac{3(T_\tau(k;\underline{\mathbf{X}}_n) - 1)}{T_\tau(k;\underline{\mathbf{X}}_n) - 3}\right). \tag{2.2}$$

Under adequate general conditions, and for an appropriate tuning parameter τ, the ρ-estimators in (2.2) show highly stable sample paths as functions of k, the number of top o.s.'s used, for a range of large k-values (see, for instance, the pattern of $\hat\rho_0(k)$ in Figure 4). It is sensible to advise practitioners not to choose blindly the value of τ in (2.2): sample paths of $\hat\rho_\tau(k)$, as functions of k, for a few values of τ, should be drawn, in order to elect the value of τ which provides higher stability for large k, by means of any stability criterion. The value $\tau = 0$, considered in the description of the algorithms, is however the most adequate whenever we are in the region $|\rho| \le 1$, a quite common region in applications and the region where bias reduction is indeed needed. Distributional properties of the estimators in (2.2) can be found in Fraga Alves *et al.* (2003). Consistency is achieved for the class of models in (1.4), for *intermediate* k-values, i.e., k-values such that (1.3) holds, and also such that $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$. Interesting alternative classes of ρ-estimators have recently been introduced in Goegebeur *et al.* (2008, 2010) and in Ciuperca and Mercadier (2010).

For the estimation of the scale second-order parameter β, and given the sample

$\underline{\mathbf{X}}_n$, we shall here consider again the estimator in Gomes and Martins (2002),

$$\hat{\beta}_{\hat{\rho}}(k) \equiv \hat{\beta}_{\hat{\rho}}(k; \underline{\mathbf{X}}_n) := \left(\frac{k}{n}\right)^{\hat{\rho}} \frac{d_{\hat{\rho}}(k)\ D_0(k) - D_{\hat{\rho}}(k)}{d_{\hat{\rho}}(k)\ D_{\hat{\rho}}(k) - D_{2\hat{\rho}}(k)}, \tag{2.3}$$

dependent on an adequate ρ-estimator, $\hat{\rho}$, and where, for any $\alpha \leq 0$,

$$d_\alpha(k) := \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha} \quad \text{and} \quad D_\alpha(k) := \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha} U_i,$$

with

$$U_i := i\left(\ln \frac{X_{n-i+1:n}}{X_{n-i:n}}\right), \qquad 1 \leq i \leq k,$$

the *scaled log-spacings* associated with the sample $\underline{\mathbf{X}}_n$. It has been advised the computation of these second-order parameters' estimators at a k-value given by

$$k_1 = \left[n^{1-\epsilon}\right], \quad \epsilon = 0.001. \tag{2.4}$$

The estimator $\hat{\rho}$, used in (2.3), is thus $\hat{\rho} = \hat{\rho}_\tau(k_1; \underline{\mathbf{X}}_n)$, with $\hat{\rho}_\tau(k; \underline{\mathbf{X}}_n)$ and k_1 given in (2.2) and (2.4), respectively. With the choice of k_1 in (2.4), we have obviously the validity of condition (1.3), and whenever $\sqrt{k_1}\ A(n/k_1) \to \infty$ (an almost irrelevant restriction, from a practical point of view), we get $\hat{\rho} - \rho := \hat{\rho}_\tau(k_1) - \rho = o_p(1/\ln n)$, a condition needed, in order not to have any increase in the asymptotic variance of the new bias-corrected Hill estimator in equation (1.5), comparatively with the one of the Hill estimator, in (1.2). Details on the distributional behaviour of the estimator in (2.3) can be found in Gomes and Martins (2002) and more recently in Gomes *et al.* (2008b) and Caeiro *et al.* (2009). Again, consistency is achieved for models in (1.4), and k values such that (1.3) holds and $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$. Alternative estimators of β can be found in Caeiro and Gomes (2006) and Gomes *et al.* (2010b). Algorithms for the estimation of (β, ρ) are provided in Gomes and Pestana (2007a,b), among others. Due to the fact that $\hat{\beta} = \hat{\beta}_{\hat{\rho}}(k_1)$ and $\hat{\rho} = \hat{\rho}_\tau(k_1)$, with $\hat{\rho}_\tau(k)$, $\hat{\beta}_{\hat{\rho}}(k)$ and k_1 given in (2.2), (2.3) and (2.4), respectively, depend on $\tau \in \mathbb{R}$, we shall use the notations $\overline{H}_\tau = \overline{H}$ and $\overline{H}_\tau^{(q)} \equiv \overline{H}^{(q)}$, if needed.

2.2 A first algorithm for the heuristic choice of k and q

As mentioned before, the first algorithm, *Algorithm* I, is quite similar to the ones in Figueiredo *et al.* (2010) and Gomes *et al.* (2010a). Essentially, we only slight change Step 7., choosing a larger value for k^*, the maximum of the set $\mathcal{K}^*$, defined in Step 6. of the algorithm, instead of the minimum, as suggested before. The reason for such a decision is just to make the choice compatible with the one considered in the algorithm proposed in this article (*Algorithm* II). The algorithm is written for the choice $\tau = 0$, in (2.2), but can obviously be used for another fixed choice of τ, as well as a data-driven choice of τ provided by any of the algorithms in Gomes and Pestana (2007a,b), among others.

Algorithm I .

1. Given an observed sample $(x_1, x_2, \ldots, x_n)$, consider, for $q = -1/n,\ 0(0.05)0.5$, the observed sample of excesses, $\underline{\mathbf{x}}_n^{(q)}$, with $\underline{\mathbf{X}}_n^{(q)}$ given in (1.6), and compute $\hat\rho \equiv \hat\rho_{0,q} = \hat\rho_0(k_1; \underline{\mathbf{x}}_n^{(q)})$ and $\hat\beta \equiv \hat\beta_{0,q} := \hat\beta_{\hat\rho_0}(k_1; \underline{\mathbf{x}}_n^{(q)})$, $\hat\rho_\tau(k)$, $\hat\beta_{\hat\rho}(k)$ and k_1 given in (2.2), (2.3) and (2.4), respectively.

2. Next compute, for $k = 1, 2, \ldots, n - [nq] - 1$, the observed values of $\overline{H}_0^{(q)}(k)$, with $\overline{H}_0^{(q)}(k) \equiv \overline{H}^{(q)}(k)$ given in (1.7) $\big($note that, as mentioned before and with $\overline{H} \equiv \overline{H}_\tau$ given in (1.5), $\overline{H}_\tau^{(-1/n)} = \overline{H}_\tau\big)$.

3. Obtain j_0, the minimum value of j, a positive integer, such that $a_k^{(q)}(j) = [\overline{H}_0^{(q)}(k) \times 10^j],\ k = 1, 2, \cdots, n - [nq] - 1$, has distinct elements (in the applications provided in Section 3, we were led to $j_0 = 1$ for all data sets considered).

4. For each q and for $k > \hat k_0^H$, with $\hat k_0^H$ given by

$$\hat k_0^H := \Big(\frac{(1-\hat\rho)^2 n^{-2\hat\rho}}{-2\hat\rho\hat\beta^2}\Big)^{1/(1-2\hat\rho)}, \tag{2.5}$$

the estimate of $k_0^H := \arg\min MSE(H(k))$ in Hall (1982), consider as an estimate of γ the equal consecutive values $\overline{H}_0^{(q)}(k)$, $k \in \mathcal{K}_q$, with $\mathcal{K}_q = [k_{min}^{(q)}, k_{max}^{(q)}]$, to which is associated the largest size $l_q := k_{max}^{(q)} - k_{min}^{(q)}$.

5. Choose next $q^* := \arg\max_q l_q$.

6. Consider all those estimates, $\overline{H}_0^{(q^*)}(k)$, $k_{min}^{(q^*)} \le k \le k_{max}^{(q^*)}$, now with an extra decimal place, i.e., compute $\overline{H}_0^{(q^*)}(k) = a_k^{(q^*)}(j_0+1)/10^{j_0+1}$. Count the frequencies associated to those estimates and obtain the mode $\eta^* \equiv \eta_{q^*}$ of these values. Let us denote $\mathcal{K}^*$ the set of k-values corresponding to those estimates.

7. Take k^* as the maximum of $\mathcal{K}^*$ and the adaptive EVI-estimate $\hat{\gamma}_1 = \overline{H}_0^{(q^*)}(k^*)$.

Approximate confidence intervals (CI's) for γ can be easily obtained on the basis of the asymptotic behaviour of the estimators under consideration. See, for instance, Gomes and Pestana (2007b), for an explanation of the results stated in the two following remarks.

Remark 2.1. *With the notation* $b_{k,n,\rho} = 1 + \beta(n/k)^\rho/(1-\rho)$, *and provided that* $\sqrt{k}\ (n/k)^\rho \to \lambda$, *finite,* $\sqrt{k}\{H(k)/\gamma - b_{k,n,\rho}\}$ *is approximately Normal*$(0,1)$. *We can then get approximate* $100(1-\alpha)\%$ CI's *for* γ, *given by*

$$\left(\frac{H(k)}{b_{k,n,\rho} + \frac{\xi_{1-\alpha/2}}{\sqrt{k}}},\ \frac{H(k)}{b_{k,n,\rho} - \frac{\xi_{1-\alpha/2}}{\sqrt{k}}}\right), \tag{2.6}$$

where ξ_p *denotes the quantile of probability* p *of a standard normal d.f. If* $\lambda = 0$, *we can replace in* (2.6) *the bias summand* $\beta(n/k)^\rho/(1-\rho)$ *by* 0, *i.e., we can take* $b_{k,n,\rho} = 1$.

Remark 2.2. *If we consider a second-order MVRB EVI-estimator, and levels* k *such that* $\sqrt{k}\ (n/k)^\rho \to \lambda$, *finite, we can also easily get even simpler approximate* $100(1-\alpha)\%$ CI's *for* γ. *On the basis of the statistic* $\overline{H}$ *in* (1.5), *or even* $\overline{H}_\tau^{(q)} \equiv$

$\overline{H}^{(q)}$, *in* (1.7), *for adequate values of* τ *and* q *and for the same* k*-levels, we get the following* $100(1-\alpha)\%$ *approximate* CI *for* γ,

$$\left(\frac{\overline{H}(k)}{1+\frac{\xi_{1-\alpha/2}}{\sqrt{k}}}, \frac{\overline{H}(k)}{1-\frac{\xi_{1-\alpha/2}}{\sqrt{k}}} \right). \tag{2.7}$$

2.3 A second algorithm for the heuristic choice of k and q

Alternatively, and in order to simplify the computational procedure of identification of "runs", we now suggest the replacement of steps from **3.** up to **7.** by:

Algorithm II.

3'. In order to detect the sign of the trend in the EVI-estimates, and for the same values of q as in Step **1.**, compute $\overline{H}_0^{(q)}([n^{0.95}]) - \overline{H}_0^{(q)}([n^{0.05}])$. We shall now suppose that such a sign is *positive*, the most common situation in practice.

4'. For $k > \hat{k}_0^H/2$, with $\hat{k}_0^H$ given in (2.5), modify the patterns of the estimates, in the following simple way:

$$\widetilde{H}_0^{(q)}(k) = \begin{cases} \overline{H}_0^{(q)}(k) & \text{if} \quad k \le \hat{k}_0^H/2 \\ \max\left(\widetilde{H}_0^{(q)}(k-1), \overline{H}_0(k)\right) & \text{if} \quad k > \hat{k}_0^H/2. \end{cases} \tag{2.8}$$

5'. For each q, compute the estimate that provides the "*largest run*" of estimates (equal consecutive estimates), say $\widetilde{H}_0^{(q)}(k_{q,1})$, with a size $m_q = k_{q,2} - k_{q,1} + 1$ $\big($this means that $\widetilde{H}_0^{(q)}(k_{q,1}) = \widetilde{H}_0^{(q)}(k_{q,1}+1) = \cdots = \widetilde{H}_0^{(q)}(k_{q,2})\big)$.

6'. Choose $q^{**} := \arg\max_q m_q$.

7'. Consider $k^{**} = k_{q^{**},2}$, and the adaptive EVI-estimate, $\hat{\gamma}_2 := \widetilde{H}_0^{(q^{**})}(k^{**})$.

Remarks

- If in Step **3'.**, $\overline{H}_0^{(q)}([n^{0.95}]) - \overline{H}_0^{(q)}([n^{0.05}])$ is negative (a negative trend in k), the operator max, in equation (2.8), should be replaced by the operator min.

- If there are negative elements in any sample, the sample size should be replaced everywhere in the algorithms by the number of positive elements in the sample.

Before the application, in Section 3, of these two algorithms to real and simulated samples, we shall proceed, in Section 2.3.1, with a small-scale simulation study of the modified EVI-estimators, in (2.8).

2.3.1 A small-scale Monte Carlo simulation of the modified EVI-estimators

On the basis of a Monte-Carlo simulation with 5000 runs, we present in Figure 1 the simulated patterns of expected values, E(•), and root mean squared errors, RMSE(•), of the following EVI-estimators: $H(k)$, $\overline{H}_0^{(q)}(k)$ and $\widetilde{H}_0^{(q)}(k)$, in (1.2), (1.7) and (2.8), respectively, for $q = 0.1$, 0.2 and a Student t_4 underlying parent. For the sake of simplicity, we denote these estimates by H, $\overline{H}_0|q$ and $\widetilde{H}_0|q$ respectively. The probability density function of a Student's t_ν-model with ν degrees of freedom, is given by

$$f_{t_\nu}(t) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi\nu}\,\Gamma(\nu/2)}\,\left(1 + t^2/\nu\right)^{-(\nu+1)/2},\ t \in \mathbb{R} \quad (\nu > 0).$$

For such a model, we get $\gamma = 1/\nu$ and $\rho = -2/\nu$.

The Student t_ν model was chosen merely as an illustration, and similar results were obtained for other simulated models.

As expected, and mainly due to the stochastic fluctuations of $\hat{k}_0^H$, in (2.5), the modified estimators, in (2.8), have often a slightly larger MSE than the original PORT-MVRB estimators, in (1.7), but the increase is not drastic, and like the PORT-MVRB, the modified PORT-MVRB estimators perform better than the Hill for all k-values. The potential loss of efficiency of the modified estimator $\widetilde{H}_0^{(q)}(k)$, in(2.8), comparatively with $\overline{H}_0^{(q)}(k)$, in (1.7), both computed at optimal levels, in the sense of minimum RMSE, denoted $RMSE_0^\bullet$, can be measured by the

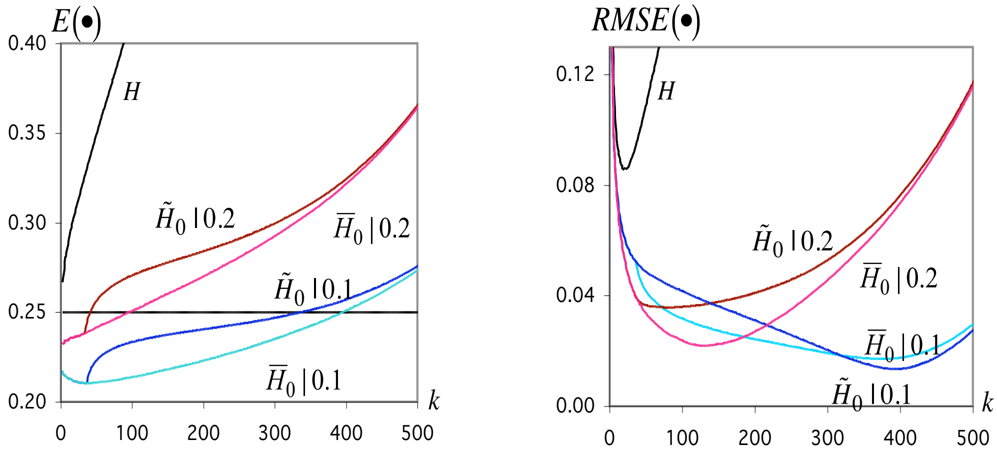

Figure 1: Comparative behaviour of the PORT-MVRB estimators in (1.7) and their modified versions in (2.8): patterns of mean values (*left*) and RMSEs (*right*), as functions of k, for an underlying Student t_4 parent ($\gamma = 0.25$ and $\rho = -0.5$).

indicator $\left(1 - RMSE_0^{\overline{H}_0}/RMSE_0^{\tilde{H}_0}\right)$, and it is never bigger than 40%. Indeed, and at optimal levels, the modified EVI-estimators can even beat the original estimators, as we could already have noticed in Figure 1, and can see in Figure 2 and Figure 3, where we picture $E_0^\bullet$, the mean values of the estimators at optimal levels, and $RMSE_0^\bullet$, respectively, as a function of the sample size n. These indicators were based on a multi-sample simulation, of 20 replicates with 5000 runs.

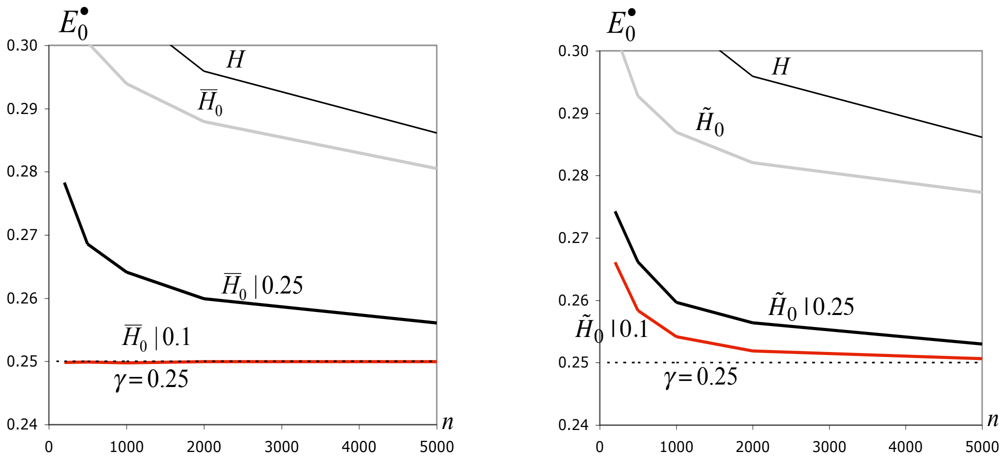

Figure 2: Mean values at optimal levels of the PORT-MVRB estimators in (1.7) (*left*) and their modified versions in (2.8) (*right*), for an underlying Student t_4 parent.

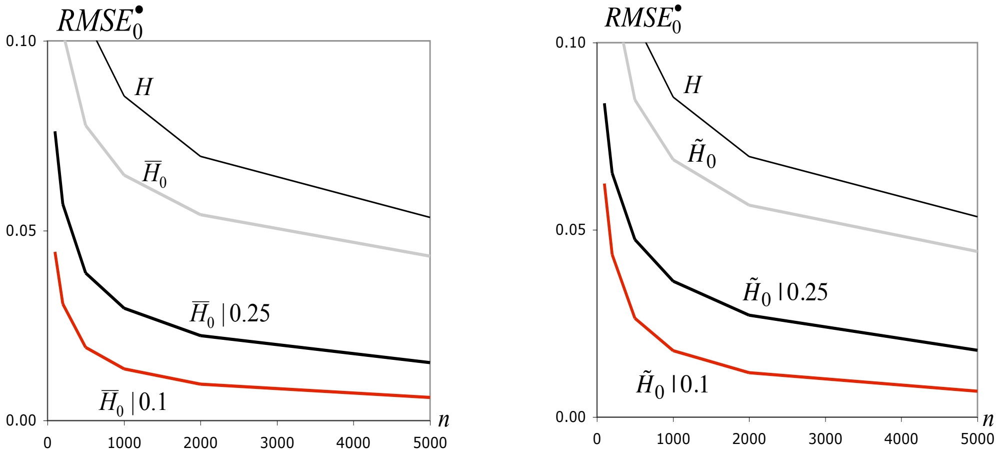

Figure 3: RMSEs at optimal levels of the PORT-MVRB estimators in (1.7) (*left*) and their modified versions in (2.8) (*right*), for an underlying Student t_4 parent.

3 Case studies and simulated samples

We shall now consider the performance of *Algorithm* I and *Algorithm* II, provided in Section 2.2 and Section 2.3, respectively, when applied to the analysis of the log-returns associated with two of the four sets of finance data considered in Gomes and Pestana (2007b). Those data sets, collected over the same period, i.e. from January 4, 1999 through November 17, 2005, were the daily closing values of the Dow Jones Industrial Average In (DJI), and Microsoft Corp. (MSFT), to be analyzed in Section 3.1 and Section 3.2, respectively. Note that the MSFT data has already been analysed, through the use of an algorithm similar to *Algorithm* I, in Gomes *et al.* (2010a). Additionally, we have considered over the same period, and in Section 3.3, the Euro-GB Pound (EGBP) daily exchange rates, also partially used in Gomes *et al.* (2008c). All these samples have a size $n = 1762$. In Section 3.4, we have dealt with the data already analysed in Drees (2003) and later in Araújo Santos *et al.* (2006), the daily log-returns of NASDAQ (NASD) index from 1997 to 2000, which corresponds to a sample size $n = 1037$. Although there is some increasing trend in the volatility of all these log-returns, stationarity and weak dependence are assumed, under the same considerations as in Drees (2003). In Section 3.5, we consider the simulated Student sample in Gomes *et al.* (2010a) and two other sequentially simulated samples, all with a size $n = 1762$ and from a Student t_ν model with $\nu = 4$ degrees of freedom.

3.1 DJI data

For this first data set, we present two figures. In the first one, Figure 4, we picture, at the *left*, a box-and-whiskers' plot of the available data. At the *right* we present, for the original sample, the sample path of the $\hat{\rho}_\tau(k)$ estimates in (2.2), as function of k, for $\tau = 0$ and $\tau = 1$, together with the sample paths of the β-estimators in (2.3), also for $\tau = 0$ and $\tau = 1$.

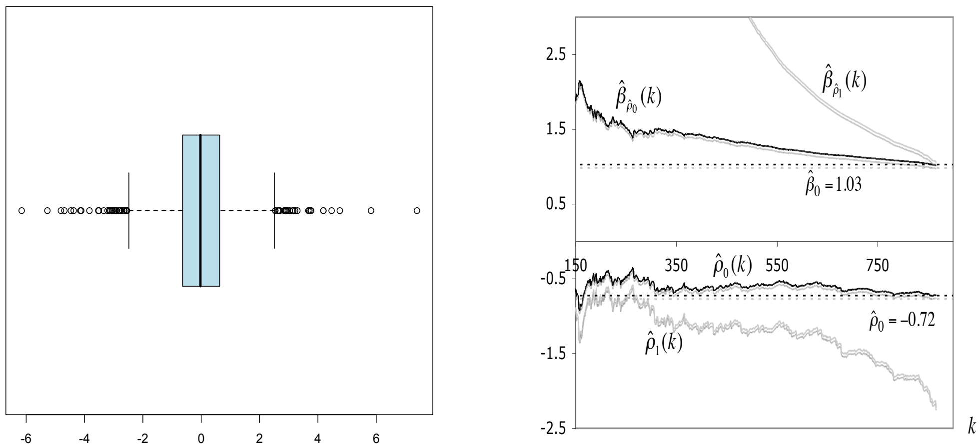

Figure 4: Box-and-whiskers plot (*left*) and estimates of the shape second-order parameter ρ and of the scale second-order parameter β (*right*) for the DJI data.

From Figure 4, *left*, we immediately see that the underlying model has heavy left and right tails. Indeed, this happens for all log-returns' data sets and it is the reason why we have eliminated in all these case-studies the estimators associated with $q = 0$, due to their inconsistency (see Gomes *et al.*, 2008a, for details). The number of positive elements in the available sample of DJI log-returns is $n_0 = 867$. We have been led to the ρ-estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.72$, obtained at the level $k_1 = [n_0^{0.999}] = 861$. The associated β-estimate is $\hat{\beta} \equiv \hat{\beta}_0 = 1.03$ (see Figure 4, *right*). Note that the sample paths of the ρ-estimates associated with $\tau = 0$ and $\tau = 1$ lead us to choose, on the basis of any stability criterion for large k, the estimate associated with $\tau = 0$. This happens not only for this sample but for all samples studied in this section.

Finally, in Figure 5, we present the adaptive and non-adaptive estimates estimates of γ, provided by H and $\overline{H}_0^{(q)}$ (*left*) and by H and $\widetilde{H}_0^{(q)}$ (*right*), $q = -1/n$, 0.1 and 0.3 ($\overline{H}_0^{(-1/n)} = \overline{H}_0 = \overline{H}$ and $\widetilde{H}_0^{(-1/n)} = \widetilde{H}_0 = \widetilde{H}$), with H, $\overline{H}_0^{(q)} \equiv \overline{H}^{(q)}$ and $\widetilde{H}_0^{(q)}$ given in (1.2), (1.7) and (2.8), respectively.

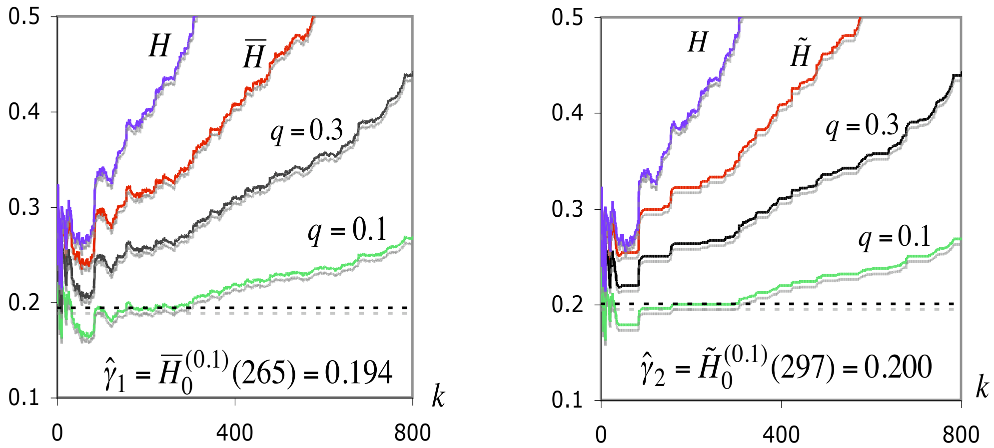

Figure 5: Non-adaptive EVI-Estimates for the DJI data and adaptive estimates obtained through *Algorithm* I (*left*) and through *Algorithm* II (*right*).

Note that the Hill estimator $H(k)$, in (1.2), is unbiased for the estimation of the extreme value index γ only when the underlying model is a strict Pareto model. Otherwise, i.e. when we have only Pareto-like tails, as surely happens here and can be seen from Figure 5, it exhibits a quite relevant bias. The PORT-MVRB estimators, $\overline{H}^{(q)}$, in (1.7), or their modified versions $\widetilde{H}^{(q)}$, in (2.8), based on shifted $MVRB$ EVI-estimators, which are "asymptotically unbiased" for adequate values of q, have a smaller bias, exhibit more stable sample paths as functions of k, and enable us to take a decision upon the estimate of γ and other parameters of extreme events to be used, with the help of any heuristic stability criterion, like the "*largest run*" method suggested in Gomes *et al.* (2005), and the ones written algorithmically in Section 2.2 and Section 2.3, regarding adaptive EVI-estimation.

- In Step **5.** of *Algorithm* I, we have been led to the choice $q^* = 0.1$, associated with a run of size 610 of a γ-estimate equal to 0.2. In Step **6.**, we have got a mode $\eta^* = 112$ for the γ-estimate 0.19, $k^* = 265$ and the adaptive PORT-MVRB estimate is $\hat{\gamma}_1 = 0.194$. The associated 99% MVRB-confidence interval for γ, based upon (2.7), is $(0.167, 0.230)$, with a size equal to 0.063.
- The largest run, in Step **5'**. of *Algorithm* II, has a size equal to 140 and was attained by the estimate $\hat{\gamma}_2$ =0.200, and values of k such that $158 \le k \le 297$. We have then been led to the choice $q^{**} = 0.1$ and $k^{**} = 297$. The associated 99% MVRB-confidence interval for γ, based again upon (2.7), is then $(0.174, 0.236)$, with a size equal to 0.061, slightly smaller than 0.063.
- The estimate of k_0^H, in (2.5), is 71 and $H(71) = 0.264$. The approximate 99% confidence interval for γ, given in (2.6), is then $(0.193, 0.350)$, with a size equal to 0.157.

3.2 MSFT data

Figure 6 is similar to Figure 5, now for the MSFT data.

- The number of positive elements in the available sample of log-returns is now $n_0 = 882$. We were led to the ρ-estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.72$, obtained at the level $k_1 = [n_0^{0.999}] = 876$. The associated β-estimate is $\hat{\beta} \equiv \hat{\beta}_0 = 1.02$.
- In Step **5.** of *Algorithm* I, we have been led to the choice $q^* = 0.1$, associated with a run of size 810 of a γ-estimate equal to 0.2. In Step **6.** of *Algorithm* I, we have got a mode $\eta^* = 402$ for the γ-estimate 0.24, $k^* = 808$ and the adaptive

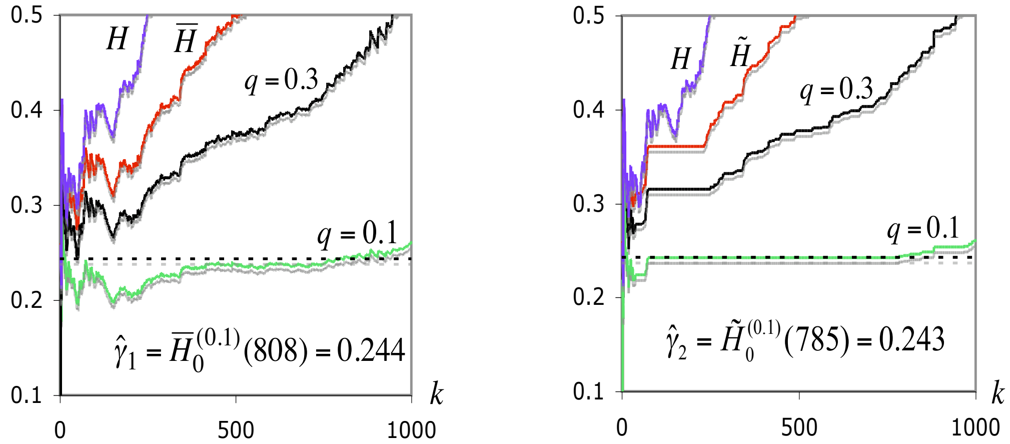

Figure 6: Non-adaptive EVI-Estimates for the MSFT data and adaptive estimates obtained through *Algorithm* I (*left*) and through *Algorithm* II (*right*).

PORT-MVRB estimate $\hat{\gamma}_1 = 0.244$. The associated 99% MVRB-confidence interval for γ is $(0.223, 0.268)$, with a size equal to 0.045.

- The largest run, in Step **5'.** of *Algorithm* II, has a size equal to 713 and was attained by the estimate $\hat{\gamma}_2$ =0.243, for $73 \le k \le 785$, as can be seen in Figure 6, *right*. We have then been led to the choice $q^{**} = 0.1$ and $k^{**} = 785$. The associated 99% MVRB-confidence interval for γ is then $(0.223, 0.268)$, with a size also equal to 0.045.

- The estimate of k_0^H, provided in (2.5), is $\hat{k}_0^H = 71$ and $H(71) = 0.391$. The approximate 99% confidence interval is then $(0.286, 0.518)$, with a size equal to 0.232, and including neither $\hat{\gamma}_1$ nor $\hat{\gamma}_2$.

3.3 Euro-GBP (EGBP) data

Figure 7 is again similar to Figure 5, now for EGBP data, and similar conclusions can be drawn.

- The number of positive elements in the available sample of log-returns is now $n_0 = 835$. We have then be led to the ρ-estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.67$, obtained at

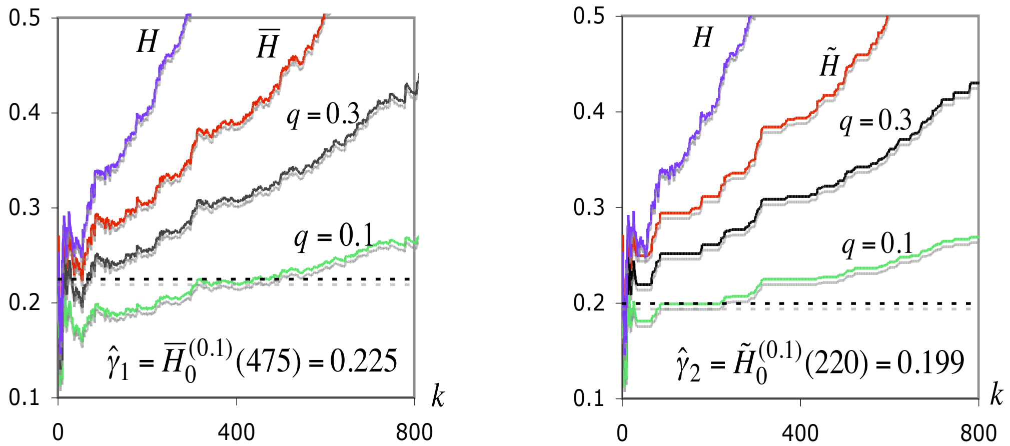

Figure 7: Non-adaptive EVI-Estimates for the EGBP data and adaptive estimates obtained through *Algorithm* I (*left*) and through *Algorithm* II (*right*).

the level $k_1 = [n_0^{0.999}] = 829$. The associated β-estimate is $\hat{\beta} \equiv \hat{\beta}_0 = 1.03$.

- In Step **5.** of *Algorithm* I, we have been led to the choice $q^* = 0.1$, associated with a run of size 603 of a γ-estimate equal to 0. In Step **6.**, we got a mode $\eta^* = 144$, $k^* = 475$, and the adaptive PORT-MVRB estimate is $\hat{\gamma}_1 = 0.225$. The associated 99% MVRB-confidence interval for γ is $(0.201, 0.255)$, with a size equal to 0.054.

- The largest run, in Step **5'**. of the Algorithm, has a size equal to 135 and was attained by the estimate $\hat{\gamma}_2$ =0.199, for $85 \leq k \leq 220$. We have then been led to the choice $q^{**} = 0.1$ and $k^{**} = 220$. The associated 99% MVRB-confidence interval for γ is then $(0.170, 0.241)$, with a size 0.071, now slightly larger than 0.054.

- The estimate $\hat{k}_0^H$, provided in (2.5), is $\hat{k}_0^H = 62$ and $H(62) = 0.281$. The approximate 99% confidence interval is then $(0.202, 0.380)$, with a size equal to 0.178, now including $\hat{\gamma}_1$, but not including $\hat{\gamma}_2$.

3.4 NASDAQ (NASD) data

Figure 8 is again similar to Figure 5, now for NASD data, and similar conclusions can be drawn. For a deeper analysis of both tails of the model underlying this sample, see Araújo Santos *et al.* (2006).

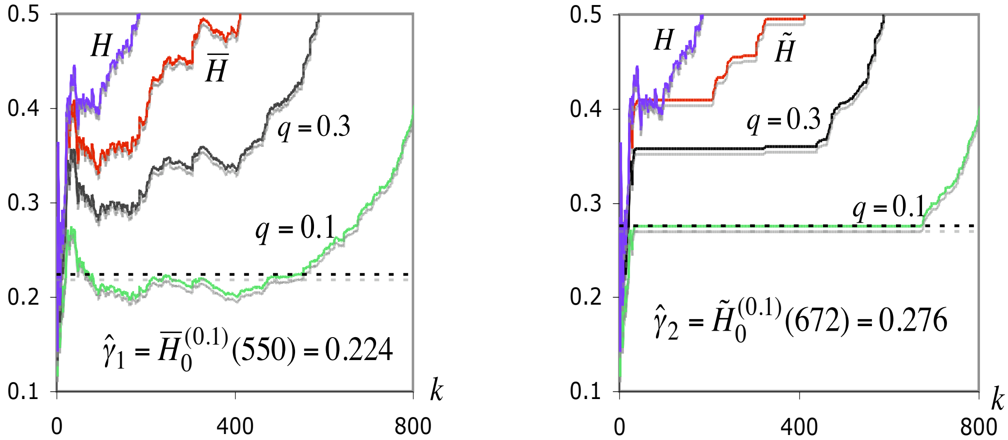

Figure 8: Non-adaptive EVI-Estimates for the NASD data and adaptive estimates obtained through *Algorithm* I (*left*) and through *Algorithm* II (*right*).

- Working with the $n_0 = 570$ positive values of the log-returns on NASD data, we got $\widehat{\rho}_0 = \widehat{\rho}_0(566) = -0.73$. The use of the β-estimate suggested before leads us to the estimate $\widehat{\beta}_0 = 1.02$.
- In Step **5.** of *Algorithm* I, we have been led to the choice $q^* = 0.1$, associated with a run of size 542 of a γ-estimate equal to 0.2. In Step **6.**, we have got a mode $\eta^* = 211$ for the γ-estimate 0.22, $k^* = 550$, and the adaptive PORT-MVRB estimate is $\hat{\gamma}_1 = 0.224$. The associated 99% MVRB-confidence interval for γ is $(0.202, 0.252)$, with a size equal to 0.050.
- The largest run, in Step **5'**. of the Algorithm, has a size equal to 641 and was attained by the estimate $\hat{\gamma}_2$ =0.276, for $32 \le k \le 672$. We have then been led to the choice $q^{**} = 0.1$ and $k^{**} = 672$. The associated 99% MVRB-confidence

interval for γ is then $(0.251, 0.306)$, with a size 0.055, again slightly larger than 0.050.

- The estimate of k_0^H, provided in (2.5), is $\hat{k}_0^H = 57$ and $H(57) = 0.400$. The approximate 99% confidence interval is then $(0.283, 0.548)$, with a size equal to 0.265, and including neither $\hat{\gamma}_1$ nor $\hat{\gamma}_2$..

3.5 Simulated Student samples

Due to the specificity of the analysed real data sets, and to the fact that log-returns have often been modelled by a Student-t or its skewed versions (see Jones and Faddy, 2003, among others), we have simulated 3 random samples of size $n = 1762$, from a Student's t_ν-model with $\nu = 4$ degrees of freedom.

- In the first sample, the number of positive elements was $n_0 = 904$. We have been led to ρ-estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.72$, obtained at the level $k_1 = [n_0^{0.999}] = 897$. The associated β-estimate was $\hat{\beta} \equiv \hat{\beta}_0 = 1.02$.
- In the second sample, the number of positive elements was $n_0 = 859$. We have been led to the ρ-estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.73$, obtained at the level $k_1 = [n_0^{0.999}] = 853$. The associated β-estimate was $\hat{\beta} \equiv \hat{\beta}_0 = 1.02$.
- In the third sample, the number of positive elements was $n_0 = 904$. We have been led to the ρ-estimate $\hat{\rho} \equiv \hat{\rho}_0 = -0.72$, obtained at the level $k_1 = [n_0^{0.999}] = 897$. The associated β-estimate was $\hat{\beta} \equiv \hat{\beta}_0 = 1.02$.

Figure 9 is equivalent to Figure 5, *left*, but for the three sequentially generated Student t_ν samples, with $\nu = 4$ degrees of freedom ($\gamma = 0.25$ and $\rho = -0.5$).

Regarding the PORT EVI-estimation, performed through *Algorithm* I, we have got the following results:

- For the first sample, and in Step **5.**, we have been led to the choice $q^* = 0.1$, associated with a run of size 702 of a γ-estimate equal to 0.2. In Step **6.**, we

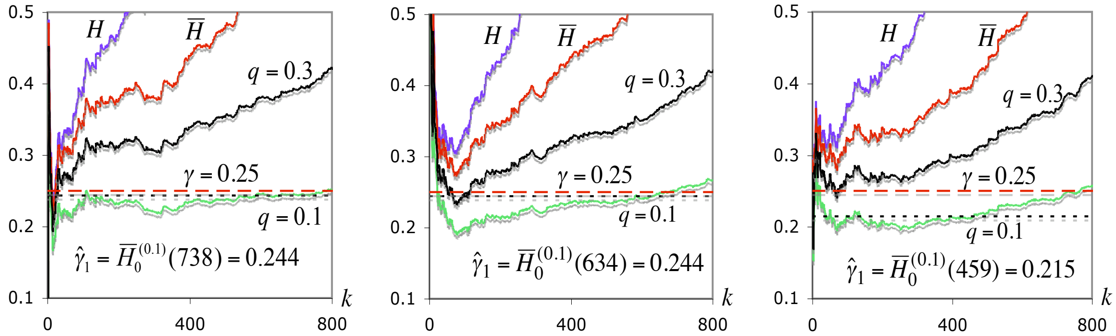

Figure 9: EVI-Estimates associated with an underlying Student t_4 model and *Algorithm* I.

have got a mode $\eta^* = 371$ for the γ-estimate 0.24, $k^* = 738$ and the adaptive PORT-MVRB estimate is $\hat{\gamma}_1 = 0.244$. The associated 99% MVRB-confidence interval for γ is $(0.222, 0.269)$, including the true value of the EVI, $\gamma = 0.25$, and with a size equal to 0.047.

- For the second sample, and in Step **5.**, we have been led to the choice $q^* = 0.1$, associated with a run of size 591 of a γ-estimate equal to 0.2. In Step **6.**, we have got a mode $\eta^* = 223$ for the γ-estimate 0.24, $k^* = 634$ and the adaptive PORT-MVRB estimate is $\hat{\gamma}_1 = 0.244$. The associated 99% MVRB-confidence interval for γ is $(0.221, 0.272)$, again including $\gamma = 0.25$, and with a size equal to 0.051.

- For the third sample, and in Step **5.**, we have been led to the choice $q^* = 0.1$, associated with a run of size 694 of a γ-estimate equal to 0.2. In Step **6.**, we have got a mode $\eta^* = 229$ for the γ-estimate 0.21, $k^* = 459$ and the adaptive PORT-MVRB estimate is $\hat{\gamma}_1 = 0.215$. The associated 99% MVRB-confidence interval for γ is $(0.192, 0.244)$, below the true value of γ, i.e., under-estimating the EVI, and with a size equal to 0.052.

- For this third sample, *Algorithm* I fails, in the sense that the 99% CI for γ does not cover the true value of γ, equal to 0.25.

Similarly, Figure 10 is equivalent to Figure 5, *right*, again for the same Student t_4 samples.

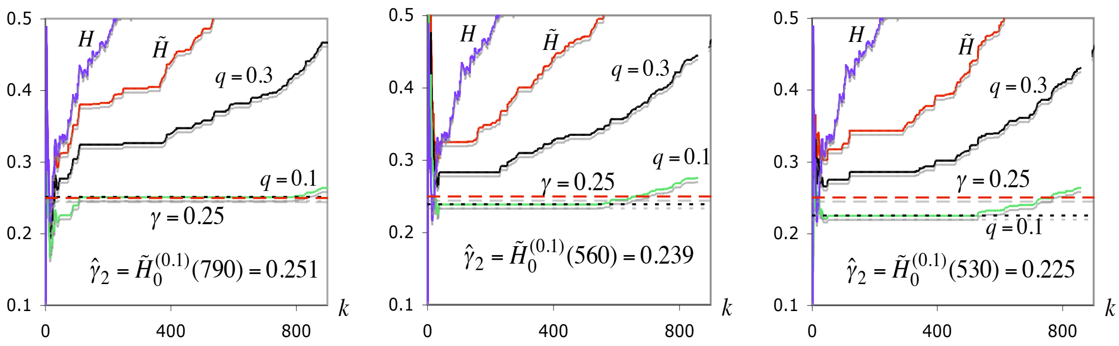

Figure 10: EVI-Estimates associated with an underlying Student t_4 model and *Algorithm* II.

Regarding the PORT EVI-estimation, performed through *Algorithm* II, we have got the following results:

- For the first sample, the largest run, in Step **5'.** of the Algorithm in Section 2.3, has a size equal to 683 and was attained by the estimate $\hat{\gamma}_2 = 0.251$, for $108 \leq k \leq 790$. We have then been led to the choice $q^{**} = 0.1$ and $k^{**} = 790$. The associated 99% MVRB-confidence interval for γ is then $(0.230, 0.276)$, with a size equal to 0.046.

- For the second sample, the largest run, in Step **5'.** of the Algorithm in Section 2.3, has a size equal to 525 and was attained by the estimate $\hat{\gamma}_2 = 0.239$, for $36 \leq k \leq 560$. We have then been led to the choice $q^{**} = 0.1$ and $k^{**} = 560$. The associated 99% MVRB-confidence interval for γ is then $(0.216, 0.269)$, with a size equal to 0.053.

- For the third sample, the largest run, in Step **5'.** of the Algorithm in Section 2.3, has a size equal to 481 and was attained by the estimate $\hat{\gamma}_2 = 0.225$, for $50 \leq k \leq 530$. We have then been led to the choice $q^{**} = 0.1$ and $k^{**} = 530$. The associated 99% MVRB-confidence interval for γ is then $(0.200, 0.256)$, with a size equal to 0.056.

- The true value of the EVI, $\gamma = 0.25$, is inside any of the three 99% CI's. The sizes of those CI's are quite similar to the ones we have got through *Algorithm* I, as expected.

Finally, regarding the adaptive Hill estimation:

- For the first sample, the estimate of k_0^H, in (2.5), was $\hat{k}_0^H = 73$ and $H(73) = 0.359$. The approximate 99% confidence interval is then $(0.257, 0.452)$, above the true value of γ, with a size equal to 0.195.

- For the second sample, we have got $\hat{k}_0^H = 71$ and $H(71) = 0.304$. The approximate 99% confidence interval is then $(0.217, 0.384)$, with a size equal to 0.167.

- For the third sample, we have got $\hat{k}_0^H = 70$ and $H(70) = 0.310$. The approximate 99% confidence interval is then $(0.216, 0.406)$, with a size equal to 0.190.

In this case we know the true value of γ, the value 0.25, and we see that such a value belongs to all 99% CI's, but the one associated with the third sample and *Algorithm* I, as well as the one associated with the first sample and the adaptive Hill's estimate. Moreover, the fact that the sizes of the MVRB-CI's are always smaller than the corresponding sizes of the Hill-CI's, even when we consider both the classical and the PORT-MVRB estimators at the same level $\hat{k}_0^H$, in (2.5), neatly favours the new methodology. The two algorithms in this article provide quite similar results, but *Algorithm* II is easier to implement not only in simulations but also for a single observed sample, and seems to provide slightly better results than *Algorithm* I, deserving thus a thorough computational Monte-Carlo study, a topic out of the scope of this article.

3.6 A general summary of the performed data analysis and final comments

In Table 1 and Table 2, we present a summary of the data analysis performed. In Table 1, apart from an indication of the sample size n, the number n_0 of positive elements in the sample, and the estimates $(\hat{\beta}, \hat{\rho})$ of the vector of second-order parameters (β, ρ), in (1.4), associated with the available original samples, we provide the values of (k, q) obtained through the implementation of the two algorithms considered, and the estimates of k for the adaptive Hill EVI-estimation, i.e., the estimates $\hat{k}_0^H$, in (2.5). In Table 2, we provide the adaptive Hill-estimates, $\hat{\gamma}^H := H(\hat{k}_0^H)$, as well as the adaptive EVI-estimates provided by *Algorithm* I, in Section 2.2, and *Algorithm* II, in Section 2.3, related with the PORT-MVRB EVI estimators discussed in this article.

Data	n	n_0	$(\hat{\beta}, \hat{\rho})$	(q^*, k^*)	(q^{**}, k^{**})	$\hat{k}_0^H$
DJI	1762	867	(1.03, −0.72)	(0.1, 265)	(0.1,297)	71
MSFT	1762	882	(1.02, −0.72)	(0.1, 808)	(0.1, 785)	71
EGBP	1762	835	(1.03, −0.67)	(0.1, 475)	(0.1,220)	62
NASD	1037	570	(1.02, −0.73)	(0.1, 550)	(0.1, 672)	57
STU1	1762	904	(1.02, −0.72)	(0.1,738)	(0.1, 790)	73
STU2	1762	859	(1.02, −0.73)	(0.1, 634)	(0.1, 560)	71
STU3	1762	904	(1.02, −0.72)	(0.1, 459)	(0.1, 530)	70

Table 1: Values of n and n_0, estimates $(\hat{\beta}, \hat{\rho})$ associated with the originl samples, and adaptive estimates of the tuning parameter q and threshold k, for the different data sets under analysis.

Overall comments

- For all data sets, the adaptive Hill EVI-estimate is even above the upper limit

Data	$\hat{\gamma}^H := H(\hat{k}_0^H)$	$\hat{\gamma}_1 := \overline{H}_0^{(q^*)}(k^*)$	$\hat{\gamma}_2 := \widetilde{H}_0^{(q^{**})}(k^{**})$
DJI	0.26 (0.193, 0.350)	0.19 (0.167, 0.230)	0.20 (0.174, 0.236)
MSFT	0.39 (0.286, 0.518)	0.24 (0.223, 0.268)	0.24 (0.223, 0.268)
EGBP	0.28 (0.202, 0.380)	0.23 (0.201, 0.255)	0.20 (0.170, 0.241)
NASD	0.40 (0.283, 0.548)	0.22 (0.202, 0.252)	0.28 (0.251, 0.306)
STU1	0.36 (0.257, 0.452)	0.24 (0.222, 0.269)	0.25 (0.230, 0.276)
STU2	0.30 (0.217, 0.384)	0.24 (0.221, 0.272)	0.24 (0.216, 0.269)
STU3	0.31 (0.216, 0.406)	0.22 (0.192, 0.244)	0.23 (0.200, 0.256)

Table 2: Adaptive EVI-estimates obtained through Hill estimators at estimated optimal level ($\hat{\gamma}^H$), *Algorithm* I ($\hat{\gamma}_1$) and *Algorithm* II ($\hat{\gamma}_2$), for the different data sets under analysis.

of the 99% CI associated with the adaptive PORT-MVRB EVI-estimates, obtained through either *Algorithm* I or *Algorithm* II.

- The other way round, the adaptive PORT-MVRB EVI-estimates are often below the lower limit of the 99% CI associated with the Hill EVI-estimate.

- For one of the generated Student samples, the first one, the true value of γ does not even belong to the 99% CI associated with the adaptive Hill EVI-estimate. It is thus clear that Hill's estimation leads to a strong over-estimation of the EVI.

- The PORT-MVRB estimation can lead in some situations to a slight under-estimation of the EVI, but the results presented clearly indicate an overall best performance of any of the data-driven methods provided. Moreover, for all simulated samples, the true value of the EVI belongs to the 99% CI associated with the adaptive estimation provided by *Algorithm* II, a point in favour of this algorithm.

- These case studies claim obviously for a comparative simulation study of *Algo-*

rithm I and *Algorithm* II, presented in Section 2.2 and Section 2.3, respectively, or at least for a thorough computational Monte-Carlo study of the properties of the adaptive estimates provided by *Algorithm* II, but these are topics out of the scope of this article.

References

[1] Araújo Santos, P., Fraga Alves, M.I. and Gomes, M.I. (2006). Peaks over random threshold methodology for tail index and quantile estimation. *Revstat* **4**:3, 227-247.

[2] Beirlant, J., Dierckx, G., Goegebeur, Y. and Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes* **2**, 177-200.

[3] Caeiro, F. and Gomes, M.I. (2006). A new class of estimators of a "scale" second order parameter. *Extremes* **9**, 193-211, 2007.

[4] Caeiro, F., Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat* **3**:2, 111-136.

[5] Caeiro, F., Gomes, M.I. and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework. *Commun. in Statist. – Theory and Methods* **38**:7, 1019-1040.

[6] Ciuperca, G. and Mercadier, C. (2010). Semi-parametric estimation for heavy tailed distributions. *Extremes* **13**:1, 55-87.

[7] Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **9**:4, 617-657.

[8] Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760-781.

[9] Figueiredo, F., Gomes, M.I., Henriques-Rodrigues, L. and Miranda, C. (2010). A Quasi-PORT Methodology for VaR Based on Second-Order Reduced-Bias Estimation. *J. Statist. Comput. and Simul.* DOI: 10.1080/00949655.2010.547196.

[10] Fraga Alves, M.I., Gomes M.I. and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica* **60**:2, 194-213.

[11] Goegebeur, Y., Beirlant, J. and de Wet, T. (2008). Linking Pareto-tail kernel goodness-of-fit statistics with tail index at optimal threshold and second order estimation. *Revstat* **6**:1, 51-69.

[12] Goegebeur, Y., Beirlant, J. and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *J. Statist. Plann. Inference* **140**, 2632-2652.

[13] Gomes, M.I. and Martins, M.J. (2002). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes* **5**:1, 5-31.

[14] Gomes, M.I. and Pestana, D. (2007a). A simple second order reduced-bias' tail index estimator. *J. Statist. Comp. and Simul.* **77**:6, 487-504.

[15] Gomes, M.I. and Pestana, D. (2007b). A sturdy reduced-bias extreme quantile (VaR) estimator. *J. American Statistical Association* **102**: 477, 280–292.

[16] Gomes, M.I., Martins, M.J. and Neves, M.M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events – the Jackknife methodology. *Extremes* **3**:3, 207-229.

[17] Gomes, M.I., Figueiredo, F. and Mendonça, S. (2005). Asymptotically best linear unbiased tail estimators under a second order regular variation condition. *J. Statist. Plann. Infer.* **134**:2, 409-433.

[18] Gomes, M.I., Martins, M.J. and Neves, M.M. (2007). Improving second order reduced bias extreme value index estimation. *Revstat* **5**:2, 177-207.

[19] Gomes, M.I., Fraga Alves, M.I. and Araújo Santos, P. (2008a). PORT Hill and Moment Estimators for Heavy-Tailed Models. *Commun. Statist. – Simulation and Comput.* **37**, 1281-1306.

[20] Gomes, M.I., de Haan, L. and Henriques-Rodrigues, L. (2008b). Tail Index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *J. Royal Statist. Soc.* **B70**: 1, 31-52.

[21] Gomes, M.I., Henriques-Rodrigues, L. and Miranda, C. (2010a). Reduced-Bias Location-Invariant Extreme Value Index Estimation: a Simulation Study. *Comm. in Statist. – Simul. and Comput.* DOI: 10.1080/03610918.2010.543297

[22] Gomes, M.I., Henriques-Rodrigues, L., Pereira, H. and Pestana, D. (2010b). Tail index and second order parameters' semi-parametric estimation based on the log-excesses. *J. Statist. Comput. and Simul.* **80**:6, 653-666.

[23] Hall, P. (1982). On some simple estimates of an exponent of regular variation. *J. Royal Statist. Soc.* **B44**, 37-42.

[24] Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163-1174.

[25] Jones, M.C. and Faddy M.J. (2003). A skew extension of the t-distribution, with applications. *J. Royal Statist. Soc.* **B 65**:1, 159-174.