Consistent estimation of a scale second-order parameter related to the PORT methodology*

Lígia Henriques-Rodrigues
CEAUL, University of Lisbon and Instituto Politécnico de Tomar

and

M. Ivette Gomes
CEAUL and DEIO, FCUL, University of Lisbon

January 15, 2013

Abstract

For heavy right tails and under a semi-parametric framework, we introduce a class of location invariant estimators of a scale second-order parameter and study its asymptotic degenerate behaviour. This class is based on the PORT methodology, with PORT standing for peaks over random thresholds. Consistency of the new class of estimators is achieved under a second-order condition on the right tail of the underlying model F and for intermediate ranks. An illustration of the finite sample behaviour of the estimators is provided through a small-scale Monte-Carlo simulation study.

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

Keywords and phrases. Asymptotic properties; location/scale invariant estimation; Monte Carlo simulation; PORT methodology; semi-parametric estimation; scale second-order parameters; statistics of extremes.

---

[*] Research partially supported by National Funds through **FCT** — Fundação para a Ciência e a Tecnologia, project PEst-OE/MAT/UI0006/2011, PTDC/FEDER, EXTREMA project and grant SFRH/BPD/77319/2011.

1 Introduction

Our interest lies in heavy right tails, i.e. we are dealing with a random sample $\underline{X}_n = (X_1, \ldots, X_n)$ from an underlying distribution function (d.f.) F with a regularly varying right tail. This means that, for a positive real γ, the right tail-function $\overline{F} := 1 - F$ is such that

$$\lim_{t\to\infty} \overline{F}(tx)/\overline{F}(t) \; = \; x^{-1/\gamma}, \quad \text{for all } x > 0, \tag{1.1}$$

i.e. $\overline{F}$ is a regularly varying function at infinity with an index of regular variation equal to $-1/\gamma$, $\gamma > 0$. We then use the notation $\overline{F} \in RV_{-1/\gamma}$.

Let us define

$$G_\gamma(x) := \begin{cases} \exp\left(-(1+\gamma\, x)^{-1/\gamma}\right), & 1+\gamma\, x > 0, \quad \text{if} \quad \gamma \neq 0 \\ \exp(-\exp(-x)), \quad x \in \mathbb{R}, & \qquad\qquad\quad\; \text{if} \quad \gamma = 0, \end{cases} \tag{1.2}$$

the so-called general extreme-value (EV) distribution. If (1.1) holds, we are in the domain of attraction for maxima of G_γ, with $\gamma > 0$, in the sense that it is possible to linearly normalise the sequence of maximum values $\{X_{n:n} := \max(X_1, \ldots, X_n)\}_{n\geq 1}$, getting convergence to a non-degenerate random variable (r.v.) with d.f. G_γ, in (1.2) (Gnedenko, 1943). We then write $F \in \mathcal{D}_{\mathcal{M}}(G_{\gamma>0})$. This type of heavy-tailed models appears often in practice, in fields like telecommunication traffic (see Resnick, 1997, and Gomes, 2003), finance, insurance, economics, ecology (see Reiss and Thomas, 2001, 2007) and biometry (see Hüsler, 2009), among others. The parameter γ, in (1.2), is the extreme-value index (EVI), one of the primary parameters of extreme events.

Let $F^{\leftarrow}$ denote the generalised inverse function of F, defined by $F^{\leftarrow}(t) := \inf\{x : F(x) \geq t\}$, and let U be the reciprocal quantile function of the r.v. X, defined as $U(t) := F^{\leftarrow}(1-1/t)$, $t \geq 1$. For heavy right tails, we assume the validity of one of the first-order conditions below:

$$F \in \mathcal{D}_{\mathcal{M}}(G_{\gamma>0}) \quad \Longleftrightarrow \quad \overline{F} \in RV_{-1/\gamma} \quad \Longleftrightarrow \quad U \in RV_\gamma. \tag{1.3}$$

The second equivalence above was proved in de Haan (1984). For several technical proofs in the field of extreme value theory we further need information about the rate of convergence in

(1.3), assuming that for every $x > 0$,

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \psi_\rho(x) := \begin{cases} \frac{x^\rho - 1}{\rho} & \text{if} \quad \rho < 0 \\ \ln x & \text{if} \quad \rho = 0, \end{cases} \tag{1.4}$$

where $|A|$ must then be in RV_ρ (Geluk and de Haan, 1987). Sometimes we also need information on the rate of convergence in (1.4), and assume that for all $x > 0$,

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} - \psi_\rho(x)}{B(t)} = \begin{cases} \frac{x^{\rho+\rho'} - 1}{\rho+\rho'} & \text{if} \quad \rho < 0 \vee \rho' < 0 \\ \ln x & \text{if} \quad \rho = \rho' = 0, \end{cases} \tag{1.5}$$

where $|B|$ must then be in $RV_{\rho'}$. For technical simplicity, we assume that $\rho < 0$ and that we can choose $A(t) = \gamma\beta t^\rho$, in (1.4), with β a non-null real number or even a slowly varying function, i.e. a regularly varying function with an index of regular variation equal to zero. This is equivalent to say that we are working with Pareto right tails such that for $C > 0$,

$$U(t) = Ct^\gamma\big(1 + \gamma\beta t^\rho/\rho + o(t^\rho)\big). \tag{1.6}$$

The pair of second-order parameters (β, ρ), in (1.6), rules the rate of convergence in (1.4) and is dependent on a possible shift in the data. More precisely, if we have a location or shift parameter $s \in \mathbb{R}$, not necessarily null, i.e. if $F(x) = F_s(x) = F_0(x - s)$, then $U(t) \equiv U_s(t) = U_0(t) + s$ and also $(\beta, \rho) = (\beta_s, \rho_s)$ depend obviously on s, with

$$(\beta_s, \rho_s) := \begin{cases} (-s/C, -\gamma) & \text{if} \quad \gamma + \rho_0 < 0 \wedge s \neq 0 \\ (\beta_0 - s/C, \rho_0) & \text{if} \quad \gamma + \rho_0 = 0 \ \wedge \ s \neq 0 \\ (\beta_0, \rho_0) & \qquad \text{otherwise}, \end{cases} \tag{1.7}$$

where β_0 and ρ_0 are respectively the scale and shape second-order parameters associated to an unshifted model ($s = 0$). Further details on the influence of a shift $s \neq 0$ in the second-order parameters are given in the Appendix.

The adequate estimation of the second-order parameters β and ρ is of primordial importance in the adaptive choice of the best number of top order statistics (o.s.'s) to be considered in the EVI-estimation, as well as in the construction of second-order reduced-bias (SORB) or minimum-variance reduced-bias (MVRB) EVI-estimators. Overviews of the subject can be found in Chapter 6 of the book by Reiss and Thomas (2007), Gomes *et al.* (2008a) and Beirlant *et al.* (2012), among others. However, despite of scale-invariant, many classes of EVI-estimators are not location-invariant, but such invariance can be attained through the use of the PORT-methodology, introduced in the sequel, with PORT standing for peaks over random thresholds, the terminology used in Araújo Santos *et al.* (2006).

Let $X_{i:n}$, $1 \le i \le n$, denote the o.s.'s associated to the random sample $\underline{X}_n = (X_1, \ldots, X_n)$ from an underlying d.f. F. The class of estimators suggested here is a function of the sample of excesses over a random threshold $X_{n_q:n}$, with $n_q = \lfloor nq \rfloor + 1$, where $\lfloor x \rfloor$ stands for the integer part of x. Such a sample is denoted by

$$\underline{X}_n^{(q)} := \left(X_{n:n} - X_{n_q:n}, X_{n-1:n} - X_{n_q:n}, \ldots, X_{n_q+1:n} - X_{n_q:n} \right), \tag{1.8}$$

where, we can have

- $0 < q < 1$, for any $F \in D_{\mathcal{M}}(G_{\gamma>0})$ (the random threshold, $X_{n_q:n}$, is an empirical quantile);
- $q = 0$, for d.f.'s with a finite left endpoint $x_{_F} := \inf\{x : F(x) > 0\}$ (the random threshold is the minimum, $X_{1:n}$).

Any statistical inference methodology based on the sample of excesses $\underline{X}_n^{(q)}$, defined in (1.8), is called a PORT-methodology. This methodology enabled the introduction and study of classical location/scale invariant EVI-estimators, like the PORT-Hill and the PORT-Moment estimators in Araújo Santos *et al.* (2006). These PORT EVI-estimators were further studied for finite-samples in Gomes *et al.* (2008b). This methodology was also applied in the estimation of high quantiles in Henriques-Rodrigues and Gomes (2009). PORT MVRB EVI-estimators have been studied for finite samples and by Monte-Carlo simulation in Gomes *et al.* (2011a, 2011b).

The PORT methodology leads to location-invariant estimation, where the unshifted model F_0 plays thus a central role. In what follows, we use the notation χ_q for the q-quantile of the d.f. F_0, i.e. the value $\chi_q := F_0^{\leftarrow}(q)$ (by convention $\chi_0 := x_{_F}$, the left endpoint of F_0). Since $n_q/n \to q$, as $n \to \infty$, we then know that the o.s. $X_{n_q:n}$, associated to a sample from F_0, is a consistent estimator for $F_0^{\leftarrow}(q)$ (van der Vaart, 1998, p.308), i.e. we have the following convergence in probability:

$$X_{n_q:n} \underset{n\to\infty}{\overset{p}{\longrightarrow}} \chi_q = F_0^{\leftarrow}(q), \quad \text{for} \quad 0 \le q < 1 \quad \left(\chi_0 = x_{_F}\right). \tag{1.9}$$

When applying the PORT-methodology, we are working with the sample of excesses in (1.8), and we can assume that we are dealing with an unshifted d.f. F_0 underlying the r.v. X_0, to which we are inducing a random shift, strictly related to χ_q, in (1.9). More precisely, we have a shift $s = -\chi_q$, i.e. we are working with $X_q := X_0 - \chi_q$, and use the simpler notation (β_q, ρ_q) for $(\beta_{-\chi_q}, \rho_{-\chi_q})$, with (β_s, ρ_s) defined in (1.7). Hence

$$(\beta_q, \rho_q) := \begin{cases} (\chi_q/C, -\gamma) & \text{if } \ \gamma + \rho_0 < 0 \wedge \chi_q \neq 0 \\ (\beta_0 + \chi_q/C, \rho_0) & \text{if } \ \gamma + \rho_0 = 0 \ \wedge \ \chi_q \neq 0 \\ (\beta_0, \rho_0) & \quad \text{otherwise.} \end{cases} \tag{1.10}$$

A class of location-invariant semi-parametric estimators of the so-called PORT-ρ second-order parameter, ρ_q, in (1.10), was recently introduced and studied in Henriques-Rodrigues and Gomes (2012) and Henriques-Rodrigues *et al.* (2012). These authors mention that the main motivation for the theoretical study of a class of estimators of the shape second-order parameter ρ_q is related to its possible use, concomitantly with a class of PORT estimators of the functional A, in (1.4), or at least of an adequate location-invariant estimator of the scale parameter of such a A-function, in the study of the asymptotic behaviour of second-order PORT-MVRB EVI-estimators, invariant for changes in location. We are now interested in the asymptotic behaviour of a class of location-invariant semi-parametric estimators of the so-called PORT-β second-order parameter, β_q, also in (1.10).

In Section 2, we introduce the new class of PORT-β estimators of the second-order parameter β_q in (1.10). In Section 3, we present a few preliminary asymptotic results related to the PORT-methodology. In Section 4 we justify the class of PORT-β estimators of the scale second-order parameter β_q, addressing the possibility of shifted heavy-tailed models, and refer the conditions required for their consistency. In Section 5, we illustrate the finite sample behaviour of the new estimators through a small-scale Monte-Carlo simulation study. In Section 6, we present the proofs of the results stated in Section 4. Finally, in the Appendix we provide further details on the influence of a shift $s \neq 0$ in the second and third-order parameters.

2 The class of semi-parametric PORT-β estimators

The building block of our estimators of the scale second-order parameter β_q, defined in (1.10) are the statistics used in Dekkers *et al.* (1989), Gomes *et al.* (2002), Fraga Alves *et al.* (2003), Caeiro and Gomes (2006), Henriques-Rodrigues and Gomes (2012) and Henriques-Rodrigues *et al.* (2012), among others, i.e. for $\alpha > 0$ we consider the moment statistics of the log-excesses,

$$M_{n,k}^{(\alpha)} \equiv M_{n,k}^{(\alpha)}(\underline{X}_n) := \frac{1}{k}\sum_{i=1}^{k}\left(\ln X_{n-i+1:n} - \ln X_{n-k:n}\right)^{\alpha}, \tag{2.1}$$

but now applied to the sample of excesses $\underline{X}_n^{(q)}$, $0 \le q < 1$, in (1.8). For an intermediate k-sequence, i.e. a sequence $k = k_n$ of positive integers such that

$$k = k_n \to \infty \quad \text{and} \quad k = o(n) \quad \text{as} \quad n \to \infty, \tag{2.2}$$

we shall thus consider the location and scale-invariant statistics,

$$M_{n,k}^{(\alpha,q)} \equiv M_{n,k}^{(\alpha)}(\underline{X}_n^{(q)}) := \frac{1}{k}\sum_{i=1}^{k}\left(\ln\frac{X_{n-i+1:n} - X_{n_q:n}}{X_{n-k:n} - X_{n_q:n}}\right)^{\alpha}, \tag{2.3}$$

defined for $k < n - n_q$, with $M_{n,k}^{(\alpha)}(\underline{X}_n) \equiv M_{n,k}^{(\alpha)}$ given in (2.1), $\alpha > 0$.

Let $\mathbb{E}$ and $\mathbb{V}ar$ denote the mean value and variance operators, respectively, let E denote a unit exponential r.v. and let $\Gamma(t)$ denote the complete Gamma function. For any real $\alpha > 0$, with $\gamma \ge 0$ and $\rho < 0$, let us define

$$\begin{aligned}
\mu_\alpha^{(1)}(\gamma) &:= \mathbb{E}\Big(E^\alpha e^{-\gamma E}\Big) = \frac{\Gamma(\alpha+1)}{(1+\gamma)^{\alpha+1}} \quad \mu_\alpha^{(1)} := \mu_\alpha^{(1)}(0) = \Gamma(\alpha+1), && (2.4)\\
\sigma_\alpha^{(1)} &:= \sqrt{\mathbb{V}\mathrm{ar}(E^\alpha)} = \sqrt{\Gamma(2\alpha+1) - \Gamma^2(\alpha+1)}, && (2.5)\\
\mu_\alpha^{(2)}(\gamma,\rho) &:= \mathbb{E}\big(E^{\alpha-1}\, e^{-\gamma E}\, (e^{\rho E}-1)/\rho\big) = \frac{\Gamma(\alpha)}{\rho}\Big(\frac{(1+\gamma)^\alpha - (1+\gamma-\rho)^\alpha}{(1+\gamma-\rho)^\alpha(1+\gamma)^\alpha}\Big), &&\\
\mu_\alpha^{(2)}(\rho) &:= \mu_\alpha^{(2)}(0,\rho) = \frac{\Gamma(\alpha)}{\rho}\Big(\frac{1-(1-\rho)^\alpha}{(1-\rho)^\alpha}\Big), &&\\
\overline{\mu}_\alpha^{(2)}(\rho) &:= \mu_\alpha^{(2)}(\rho)/\mu_\alpha^{(1)}, \qquad \overline{\sigma}_\alpha^{(1)} := \sigma_\alpha^{(1)}/\mu_\alpha^{(1)}, && (2.6)
\end{aligned}$$

and for any $\theta_1,\ \theta_2 > 0$,

$$d_{\alpha,\theta_1,\theta_2}(\rho) \;:=\; \overline{\mu}_{\alpha\theta_1}^{(2)}(\rho) - \overline{\mu}_{\alpha\theta_2}^{(2)}(\rho). \tag{2.7}$$

For *tuning* parameters $\eta_q \in \mathbb{R}$ (as detected in Caeiro and Gomes, 2006), $\alpha, \theta_1, \theta_2 \in \mathbb{R}^+$, $\theta_1 \neq \theta_2$, we shall consider the PORT-versions of the r.v.'s used in the aforementioned paper for the estimation of β, in (1.6), i.e.

$$D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma) := \left(\frac{M_{n,k}^{(\alpha\theta_1,q)}}{\mu_{\alpha\theta_1}^{(1)}}\right)^{\eta_q/\theta_1} - \left(\frac{M_{n,k}^{(\alpha\theta_2,q)}}{\mu_{\alpha\theta_2}^{(1)}}\right)^{\eta_q/\theta_2}, \tag{2.8}$$

with $M_{n,k}^{(\alpha,q)}$ and $\mu_\alpha^{(1)}$ defined in (2.3) and (2.4), respectively. As detailed in Section 4, under adequate conditions upon the growth of $k = k_n$, the study of the asymptotic behaviour of the r.v.'s $D_{n,k}^{\alpha,\theta_1,\theta_2,\eta_q,q}(\gamma)$, in (2.8), enables us to introduce the class of consistent β_q-estimators, invariant for changes in location, and named PORT-β, given by

$$\widehat{\beta}_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\widehat{\rho}^{(q)}) := \frac{2d_{2\alpha,\theta_1,\theta_2}(\widehat{\rho}^{(q)})}{\alpha\eta_q d_{\alpha,\theta_1,\theta_2}^2(\widehat{\rho}^{(q)})}\left(\frac{k}{n}\right)^{\widehat{\rho}^{(q)}} \frac{\left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^2}{D_{n,k}^{(2\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)}, \tag{2.9}$$

with tuning parameters $\alpha, \theta_1, \theta_2 > 0$, $\theta_1 \neq \theta_2$, $\eta_q \in \mathbb{R}$, $q \in [0,1)$, $d_{\alpha,\theta_1,\theta_2}(\rho)$ and $D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)$ given in (2.7) and (2.8), respectively, and with $\widehat{\rho}^{(q)}$ the class of consistent ρ_q-estimators, invariant for changes in location, introduced in Henriques-Rodrigues and Gomes (2012) and further studied in Henriques-Rodrigues *et al.* (2012). The class of PORT-ρ estimators of the shape second-order parameter ρ_q, similar to the simplest class of ρ-estimators in Fraga Alves *et al.* (2003), is also dependent on a tuning parameter $\tau_q \in \mathbb{R}$ and is given by

$$\widehat{\rho}_k^{(q)} \equiv \widehat{\rho}_k^{(\tau_q,q)} \equiv \widehat{\rho}_{n,k|T}^{(1,2,3,\tau_q,q)} := \frac{3\left(T_{n,k}^{(1,2,3,\tau_q,q)} - 1\right)}{T_{n,k}^{(1,2,3,\tau_q,q)} - 3} \mathbb{1}\left\{T_{n,k}^{(1,2,3,\tau_q,q)} \in (1,3)\right\}, \tag{2.10}$$

where $\mathbb{1}\{A\}$ denotes the indicator function of the event A, and with $M_{n,k}^{(\alpha,q)}$ given in (2.3),

$$T_{n,k}^{(1,2,3,\tau_q,q)} := \frac{\left(M_{n,k}^{(1,q)}\right)^{\tau_q} - \left(M_{n,k}^{(2,q)}/2\right)^{\tau_q/2}}{\left(M_{n,k}^{(2,q)}/2\right)^{\tau_q/2} - \left(M_{n,k}^{(3,q)}/6\right)^{\tau_q/3}},$$

for any $\tau_q \in \mathbb{R}$, with the notation $a^{b\tau} = b \ln a$ whenever we consider $\tau_q = 0$. Moreover, $\widehat{\rho}^{(q)} := \widehat{\rho}^{(q)}_{\lfloor n-n_q \rfloor^{0.999}}$.

Caeiro and Gomes (2006) suggest in practice the consideration of $(\alpha, \theta_1, \theta_2) = (1, 1, 2)$, for the classic β-estimation, in order to get a class of estimators dependent only on a tuning parameter $\eta \in \mathbb{R}$. Taking into account this suggestion we are led to the following functional expression for the PORT-β estimators

$$\begin{aligned}\widehat{\beta}_k^{(\eta_q,q)} &\equiv \widehat{\beta}_{n,k}^{(1,1,2,\eta_q,q)}(\widehat{\rho}^{(q)}) \\ &:= \begin{cases} -\frac{2(2-\widehat{\rho}^{(q)})^2}{\eta_q\widehat{\rho}^{(q)}} \left(\frac{k}{n}\right)^{\widehat{\rho}^{(q)}} \frac{\left[\left(M_{n,k}^{(1,q)}\right)^{\eta_q} - \left(M_{n,k}^{(2,q)}/2\right)^{\eta_q/2}\right]^2}{\left(M_{n,k}^{(2,q)}/2\right)^{\eta_q} - \left(M_{n,k}^{(4,q)}/24\right)^{\eta_q/2}} & \text{if} \quad \eta_q \neq 0 \\ -\frac{2(2-\widehat{\rho}^{(q)})^2}{\widehat{\rho}^{(q)}} \left(\frac{k}{n}\right)^{\widehat{\rho}^{(q)}} \frac{\left[\ln\left(M_{n,k}^{(1,q)}\right) - \frac{1}{2}\ln\left(M_{n,k}^{(2,q)}/2\right)\right]^2}{\ln\left(M_{n,k}^{(2,q)}/2\right) - \frac{1}{2}\ln\left(M_{n,k}^{(4,q)}/24\right)} & \text{if} \quad \eta_q = 0. \end{cases}\end{aligned} \tag{2.11}$$

This new class of PORT-β estimators depends on the tuning parameters $\eta_q \in \mathbb{R}$ and $q \in [0, 1)$, related to the PORT-methodology. These two tuning parameters provide an adequate flexible class of estimators of β_q, and their non-PORT versions, with a unique parameter, say $\eta \in \mathbb{R}$, have revealed to be suitable for practical purposes, despite of high volatile for small up to moderate k comparatively to the β-estimators in Gomes and Martins (2002). The choice of the tuning parameter q can be performed with a generalisation of the algorithm proposed in Gomes and Henriques-Rodrigues (2012). This research, is however, out of the scope of this paper.

3 Technical results related to the PORT-methodology

3.1 The second-order PORT-framework for heavy-tailed models

Under the aforementioned set-up in Section 1, the transformed r.v., $X_q = X_0 - \chi_q$, has an associated quantile function given by $U_q(t) = U_0(t) - \chi_q$. The second-order condition in (1.4) translates as

$$\lim_{t\to\infty} \frac{\ln U_q(tx) - \ln U_q(t) - \gamma \ln x}{A_q(t)} = \begin{cases} \frac{x^{\rho_q}-1}{\rho_q} & \text{if} \quad \rho_q < 0 \\ \ln x & \text{if} \quad \rho_q = 0, \end{cases} \tag{3.1}$$

for all $x > 0$. Moreover, $|A_q| \in RV_{\rho_q}$, $\rho_q \leq 0$, and A_q relates to A_0 according to the following lemma, a straightforward generalisation of Lemma 3.1, in Henriques-Rodrigues *et al.* (2012).

Lemma 3.1. *Assume $U_0 \in RV_\gamma$ satisfies the second-order condition in* (1.4) *with $\rho = \rho_0$ and $A = A_0$. Then $U_q(t) = U_0(t) - \chi_q$, with χ_q defined in* (1.9)*, is such that $U_q \in RV_\gamma$ and* (3.1) *holds with* (β_q, ρ_q) *given in* (1.10) *and*

$$A_q(t) := \begin{cases} \gamma\chi_q/U_0(t), & \textit{if} \quad \gamma+\rho_0<0 \ \wedge \ \chi_q \neq 0 \\ A_0(t)+\gamma\chi_q/U_0(t), & \textit{if} \quad \gamma+\rho_0=0 \ \wedge \ \chi_q \neq 0 \\ A_0(t), & \textit{if} \quad \gamma+\rho_0>0 \vee \chi_q=0. \end{cases} \tag{3.2}$$

3.2 Second-order framework and asymptotic behaviour of auxiliary statistics

Let us further introduce the following notations. With E_i independent and identically distributed (i.i.d.) unit exponential r.v.'s, let us define

$$Z_k^{(\alpha)} := \sqrt{k}\Big(\frac{1}{k}\sum_{i=1}^{k} E_i^\alpha - \Gamma(\alpha+1)\Big)/\sigma_\alpha^{(1)}, \tag{3.3}$$

asymptotically standard normal r.v.'s, and

$$W_k^{(\alpha,\theta_1,\theta_2)} := \overline{\sigma}_{\alpha\theta_1}^{(1)} Z_k^{(\alpha\theta_1)}/\theta_1 - \overline{\sigma}_{\alpha\theta_2}^{(1)} Z_k^{(\alpha\theta_2)}/\theta_2, \tag{3.4}$$

with $\sigma_\alpha^{(1)}$ and $\overline{\sigma}_\alpha^{(1)}$ given in (2.5) and (2.6), respectively.

We next present the asymptotic behaviour, as $n \to \infty$, of $M_{n,k}^{(\alpha,q)}$ and $D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}$, in (2.3) and (2.8), respectively, based on the sample of excesses $\underline{\mathbf{X}}_n^{(q)}$, $0 \le q < 1$, in (1.8), and obviously independent on any shift s imposed to the data. We can thus assume that $s=0$. These result were stated and proven in Henriques-Rodrigues and Gomes (2012):

Proposition 3.1 (Henriques-Rodrigues and Gomes, 2012)**.** *For intermediate k, as in* (2.2), *let us assume the validity of the second-order condition in* (1.4). *We then get for* $M_{n,k}^{(\alpha,q)}$, *in* (2.3), $\alpha>0$, $k<n-n_q$, *with* $\chi_q \neq 0$, $\mu_\alpha^{(1)}$, $\big(\overline{\mu}_\alpha^{(2)}(\rho), \overline{\sigma}_\alpha^{(1)}\big)$ *and* $Z_k^{(\alpha)}$ *defined in* (1.9), (2.4), (2.6) *and* (3.3), *respectively, the distributional representation,*

$$M_{n,k}^{(\alpha,q)} \overset{d}{=} \gamma^\alpha \mu_\alpha^{(1)}\Big\{1+\overline{\sigma}_\alpha^{(1)}\frac{Z_k^{(\alpha)}}{\sqrt{k}} + \Big(\frac{\alpha}{\gamma}\,\overline{\mu}_\alpha^{(2)}(\rho_0)A_0(n/k) + \frac{\alpha\ \chi_q\ \overline{\mu}_\alpha^{(2)}(-\gamma)}{U_0(n/k)}\Big)(1+o_p(1))\Big\}.$$

Proposition 3.2 (Henriques-Rodrigues and Gomes, 2012)**.** *For intermediate k, as in* (2.2), *let us assume the validity of the second-order condition in* (1.4). *We then get for* $D_{n,k}^{(\alpha,q)}$, *in*

(2.8), $\alpha > 0$, $k < n - n_q$, with $\chi_q \neq 0$ and $d_{\alpha,\theta_1,\theta_2}(\rho)$ given in (1.9) and (2.7), respectively, the distributional representation,

$$D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma) \overset{d}{=} \frac{\tau_q}{\sqrt{k}}\, W_k^{(\alpha,\theta_1,\theta_2)} + \Big(\frac{\alpha\eta_q d_{\alpha,\theta_1,\theta_2}(\rho_0)A_0(n/k)}{\gamma} + \frac{\alpha\eta_q\chi_q d_{\alpha,\theta_1,\theta_2}(-\gamma)}{U_0(n/k)}\Big)(1+o_p(1)), \quad (3.5)$$

where $W_k^{(\alpha,\theta_1,\theta_2)}$ is the asymptotic standard normal r.v. in (3.4).

4 Consistency of the PORT-β estimators

From the definition of the parameter β_q, in (1.10), we can see that the consistency of the PORT-β estimators is related with the vector (γ, ρ_0) of the unshifted model F_0 associated with the available data. Therefore we shall consider three different regions:

(i) $\mathcal{R}_1 := \{\gamma + \rho_0 < 0 \wedge \chi_q \neq 0\}$,

(ii) $\mathcal{R}_2 := \{\gamma + \rho_0 > 0 \vee (\gamma + \rho_0 \leq 0 \wedge \chi_q = 0)\}$,

(iii) $\mathcal{R}_3 := \{\gamma + \rho_0 = 0 \wedge \chi_q \neq 0\}$.

We now state the main result in this paper.

Theorem 4.1. *Under the validity of the second-order condition in (1.4), with $\rho < 0$, (β_q, ρ_q) defined in (1.10), $\widehat{\rho}^{(q)}$ any consistent estimator of ρ_q such that $(\widehat{\rho}^{(q)} - \rho_q)\ln(n/k) = o_p(1)$, and with $\widehat{\beta}_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}$ defined in (2.9),*

$$\widehat{\beta}_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)} \underset{n\to\infty}{\overset{p}{\longrightarrow}} \beta_q,$$

for any real $\alpha > 0$, $\eta_q \in \mathbb{R}$, $\theta_1, \theta_2 \in \mathbb{R}^+\backslash\{1\}, \theta_1 \neq \theta_2$ and $0 < q < 1$ or $q = 0$ if χ_0 is finite, provided that k is an intermediate sequence, as in (2.2), and moreover

$$\sqrt{k}/A_q(n/k) \to \infty, \text{ as } n \to \infty, \quad (4.1)$$

with $A_q(\cdot)$ defined in (3.2).

Remark 4.1. *Note that when we consider models $F_0 \in \mathcal{R}_1$, $A_0(t) = o(1/U_0(t))$ and with $A_q(t) = \gamma\chi_q/U_0(t)$, by (3.2), condition (4.1) corresponds to $\sqrt{k}/U_0(n/k) \to \infty$, as $n \to \infty$.*

For models $F_0 \in \mathcal{R}_2$*,* $1/U_0(t) = o(A_0(t))$ *and since* $A_q(t) = A_0(t)$*, condition* (4.1) *is equivalent to* $\sqrt{k}A_0(n/k) \to \infty$*, as* $n \to \infty$*. Finally, for models* $F_0 \in \mathcal{R}_3$*,* $1/U_0(t) = O(A_0(t))$ *and since* $A_q(t) = A_0(t) + \gamma\chi_q/U_0(t)$*, condition* (4.1) *is equivalent to* $\sqrt{k}A_0(n/k) \to \infty$ *or* $\sqrt{k}/U_0(n/k) \to \infty$*, as* $n \to \infty$*.*

5 A small-scale Monte-Carlo simulation

As an illustration, we next present in Figures 1 and 2, respectively the mean values (E) and the root mean square errors (RMSE), of the classical estimator, denoted by $\widehat{\beta}_k^{(\eta)}$, $\eta = -0.25$, and the PORT-β estimators $\left\{\widehat{\beta}_k^{(\eta_q,q)} \equiv \widehat{\beta}_{n,k}^{(1,1,2,\eta_q,q)}(\widehat{\rho}^{(q)})\right\}_{q=0,\ 0.1,\ 0.25}$, as defined in (2.11), for the value $\eta_q = \eta = -0.25$, as a function of the sample fraction k/n, for a sample size $n = 5000$. The results are associated to the output of a small-scale simulation of size 1000, related to underlying Fréchet parents $F_0(x) = \exp(-x^{-1/\gamma}),\ x > 0$, with $\gamma = 0.25$, and the shifted model $F_s(x) = \exp\left(-(x-s)^{-1/\gamma}\right),\ x > s$, with $s = 1$. We have here used $\tau_q = 0$, the value suggested in several other research papers for the PORT-ρ estimators $\widehat{\rho}_k^{(q)} \equiv \widehat{\rho}_k^{(\tau_q,q)}$, in (2.10). The choice $\tau = 0$ has been heuristically suggested and used before for the classic ρ-estimation in the region $|\rho| \le 1$ (see Fraga Alves *et al.* (2003), for further details). Note that for this model $\beta_0 = 0.5$, $\rho_0 = -1$, i.e. $\gamma + \rho_0 < 0$, and consequently with $C = 1$ (see Remark 6.3),

$$\beta_q = \begin{cases} \chi_q = (-\ln q)^{-\gamma} & \text{if} \quad \chi_q \neq 0 \quad (0 < q < 1) \\ \beta_0 = 0.5 & \text{if} \quad \chi_q = 0 \quad (q = 0). \end{cases}$$

Due to the symmetry of the sample paths, as shown in Figure 1, it is questionable whether we should proceed to the estimation either for small or for large values of k/n. Indeed, for large k-values, we achieve the smallest RMSE, but such k-values can be theoretically questionable and the RMSE pattern is rather strange, as can be seen in Figure 2 (right).

We now would like to emphasise the following points:

- There is only a light improvement in all estimators as the sample size increases, and a high volatility of the classical β-estimators for shifted models, as can be seen, in either Figure 1 or 2, where the RMSE of such estimator is out of range, being always above 1.5.

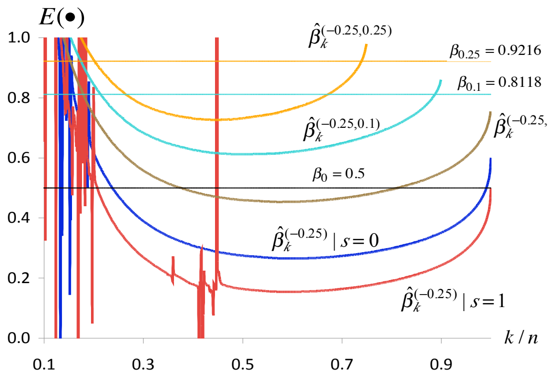

Figure 1: Mean values of the estimators under consideration for Fréchet unshifted ($s = 0$) and shifted ($s = 1$) parents, with $\gamma = 0.25$, and sample size $n = 5000$

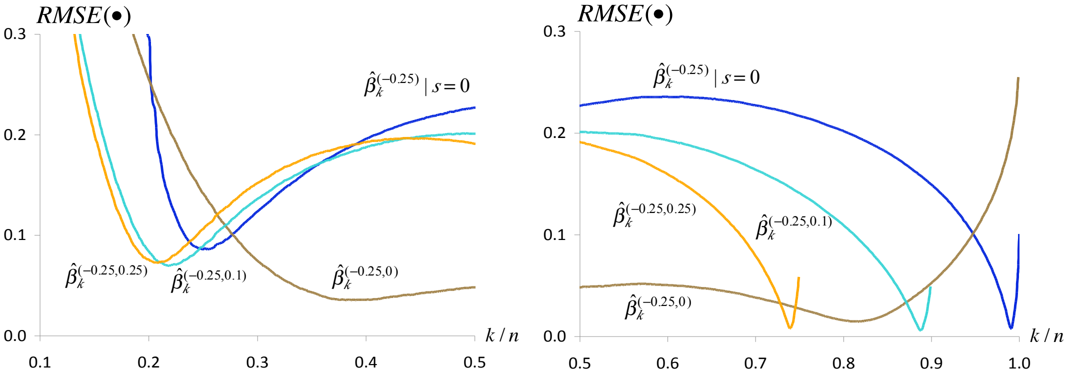

Figure 2: RMSEs of the estimators under consideration for Fréchet unshifted ($s = 0$) and shifted ($s = 1$) parents, with $\gamma = 0.25$, sample size $n = 5000$, $k/n \in (0, 0.5)$ (*left*) and $k/n \in (0.5, 1)$ (*right*)

For smaller sample sizes n, the sample paths of all estimators for small up to moderate k-values are even more volatile.

- We are in the region $\gamma + \rho_0 < 0$ ($\gamma = 0.25$, $\rho_0 = -1$, $\beta_0 = 0.5$). Consequently, the PORT-β estimator should converge to $-\chi_q = (-\ln q)^{-\gamma}$ for $\chi_q \neq 0$ and to $\beta_0 = 0.5$ for $\chi_q = 0$. The pattern of the PORT-β estimators does depend on χ_q, contrarily to the one of the PORT-ρ estimators in Henriques-Rodrigues *et al.* (2012). If we decide for a large value of k, we obtain a value close to $(-\ln q)^{-0.25}$ if $\chi_q \neq 0$, for $q = 0.1, 0.25$, but a value quite a long way from 0.5 when $\chi_q = 0$. But if we look at the region of k/n from 0.3 up to 0.8, the PORT-β estimators associated to $\chi_q = 0$ are reasonably close to $\beta = 0.5$ and quite stable, with a small RMSE.

- The PORT-β estimators associated to $\eta_q = -0.25$ are able to beat the classical one regarding minimum RMSE, even for very large sample sizes, and when we look at moderate values of k/n.
- The choice of the *tuning* parameters η and η_q is again crucial. The choice we have used here is possibly not the most adequate choice for the PORT-β estimation. This is another interesting topic out of the scope of this paper.

6 Proof of Theorem 4.1

Proof. [**Theorem 4.1**] (i) In the region $\mathcal{R}_1$, $A_0(t) = o(1/U_0(t))$, as $t \to \infty$, the last term of the right-hand side of (3.5) is the dominant one, and the r.v. $D_{n,k}^{(\alpha,\theta_1,\theta_2,\tau_q,q)}(\gamma)/(1/U_0(n/k))$ converges in probability to $\alpha\eta_q\chi_q\ d_{\alpha,\theta_1,\theta_2}(-\gamma)$ provided that (4.1) holds. Considering the first-order approximation of the function $U(t)$, in (1.6), $U_0(n/k) \equiv C(n/k)^\gamma$, we then get

$$\left(\frac{n}{k}\right)^{\gamma} D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma) \underset{n\to\infty}{\overset{p}{\longrightarrow}} \alpha\eta_q \left(\frac{\chi_q}{C}\right) d_{\alpha,\theta_1,\theta_2}(-\gamma),$$

i.e. for any $r \in \mathbb{N}$

$$\left(\frac{n}{k}\right)^{r\gamma} \left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^r \underset{n\to\infty}{\overset{p}{\longrightarrow}} \left(\alpha\eta_q \left(\frac{\chi_q}{C}\right)\right)^r d^r_{\alpha,\theta_1,\theta_2}(-\gamma). \tag{6.1}$$

To get rid of the unknown γ in $\left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^r$ it is enough to consider that

$$\left(\frac{n}{k}\right)^{\gamma} D_{n,k}^{(r\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma) \underset{n\to\infty}{\overset{p}{\longrightarrow}} r\alpha\eta_q \left(\frac{\chi_q}{C}\right) d_{r\alpha,\theta_1,\theta_2}(-\gamma). \tag{6.2}$$

The quotient between (6.1) and (6.2), enables us to say that

$$\left(\frac{n}{k}\right)^{\gamma(r-1)} \frac{\left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^r}{D_{n,k}^{(r\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)} \underset{n\to\infty}{\overset{p}{\longrightarrow}} \frac{(\alpha\eta_q\,(\chi_q/C))^{r-1}}{r} \frac{\left(d_{\alpha,\theta_1,\theta_2}\right)^r(-\gamma)}{d_{r\alpha,\theta_1,\theta_2}(-\gamma)}.$$

If we choose $r = 2$, as suggested in Caeiro and Gomes (2006), we obtain

$$\left(\frac{n}{k}\right)^{\gamma} \frac{\left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^2}{D_{n,k}^{(2\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)} \xrightarrow[n\to\infty]{p} \frac{\alpha\eta_q}{2}\left(\frac{\chi_q}{C}\right)\frac{d^2_{\alpha,\theta_1,\theta_2}(-\gamma)}{d_{2\alpha,\theta_1,\theta_2}(-\gamma)}.$$

Since, in $\mathcal{R}_1$, $\beta_q = \chi_q/C$ and $\rho_q = -\gamma$ with (β_q, ρ_q) defined in (1.10), the class of consistent r.v.'s, that converge in probability towards β_q for any $\alpha > 0$, $\eta_q \in \mathbb{R}$, $\theta_1, \theta_2 \neq \theta_2$, $0 \leq q < 1$ is given by

$$\widehat{\beta}_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\rho_q) := \frac{2d_{2\alpha,\theta_1,\theta_2}(\rho_q)}{\alpha\eta_q d^2_{\alpha,\theta_1,\theta_2}(\rho_q)}\left(\frac{n}{k}\right)^{-\rho_q} \frac{\left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^2}{D_{n,k}^{(2\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)}. \tag{6.3}$$

(ii) In the region $\gamma + \rho_0 > 0$, where $1/U_0(t) = o(A_0(t))$, as $t \to \infty$, or more generally in the region $\mathcal{R}_2$, the second term of the right-hand side of (3.5) is the dominant one, i.e. $D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)/A_0(n/k)$ converges in probability to $\alpha\eta_q d_{\alpha,\theta_1,\theta_2}(\rho_0)/\gamma$ provided that (4.1) holds. Since we can choose $A_0(t) = \gamma\beta_0 t^{\rho_0}$,

$$\left(\frac{n}{k}\right)^{-\rho_0} D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma) \xrightarrow[n\to\infty]{p} \beta_0\alpha\eta_q d_{\alpha,\theta_1,\theta_2}(\rho_0),$$

i.e., with $r \in \mathbb{N}$

$$\left(\frac{n}{k}\right)^{-r\rho} \left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^r \xrightarrow[n\to\infty]{p} (\beta\alpha\eta_q)^r\left(d_{\alpha,\theta_1,\theta_2}\right)^r(\rho). \tag{6.4}$$

Using the same type of arguments, we can get rid of the unknown γ in $\left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^r$ if we consider that

$$\left(\frac{n}{k}\right)^{-\rho_0} D_{n,k}^{(r\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma) \xrightarrow[n\to\infty]{p} \beta_0 r\alpha\eta_q d_{r\alpha,\theta_1,\theta_2}(\rho_0). \tag{6.5}$$

The quotient between (6.4) and (6.5) enables us to say that

$$\left(\frac{n}{k}\right)^{-\rho(r-1)} \frac{\left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^r}{D_{n,k}^{(r\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)} \xrightarrow[n\to\infty]{p} \frac{(\beta\alpha\eta_q)^{r-1}}{r}\frac{\left(d_{\alpha,\theta_1,\theta_2}\right)^r(\rho)}{d_{r\alpha,\theta_1,\theta_2}(\rho)}.$$

Choosing $r = 2$, as in Caeiro and Gomes (2006), and with $\beta_q = \beta_0$ and $\rho_q = \rho_0$, (β_q, ρ_q) given in (1.10), we get (6.3), i.e. a class of r.v.'s converging in probability to β_q for $\alpha > 0$, $\eta_q \in \mathbb{R}$, $\theta_1 \neq \theta_2$ and $0 \leq q < 1$.

(iii) In the region $\mathcal{R}_3$, $A_0(t)$ and $1/U_0(t)$ are of the same order, i.e., the dominant terms of the right-hand side of (3.5) are the second and the last. If we assume that (4.1) holds,

$$\frac{D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)}{A_0(n/k)} \underset{n\to\infty}{\overset{p}{\longrightarrow}} \frac{\alpha\eta_q}{\gamma}\left[d_{\alpha,\theta_1,\theta_2}(\rho_0)+\frac{\chi_q}{\beta_0 C}d_{\alpha,\theta_1,\theta_2}(-\gamma)\right].$$

Since $A_0(t)=\gamma\beta_0 t^{\rho_0}$ we then get

$$\left(\frac{n}{k}\right)^{-\rho_0} D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma) \underset{n\to\infty}{\overset{p}{\longrightarrow}} \alpha\beta_0\eta_q\left(d_{\alpha,\theta_1,\theta_2}(\rho_0)+\frac{\chi_q}{\beta_0 C}d_{\alpha,\theta_1,\theta_2}(-\gamma)\right).$$

But in $\mathcal{R}_3$, $\rho_0=-\gamma$, thence

$$\left(\frac{n}{k}\right)^{-\rho_0} D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma) \underset{n\to\infty}{\overset{p}{\longrightarrow}} \alpha\eta_q\left(\beta_0+\frac{\chi_q}{C}\right)d_{\alpha,\theta_1,\theta_2}(\rho_0).$$

Considering the same type of procedures used in cases (i) and (ii) and with $r=2$ we are led to

$$\left(\frac{n}{k}\right)^{-\rho_0}\frac{\left(D_{n,k}^{(\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)\right)^2}{D_{n,k}^{(2\alpha,\theta_1,\theta_2,\eta_q,q)}(\gamma)} \underset{n\to\infty}{\overset{p}{\longrightarrow}} \frac{\alpha\eta_q}{2}\left(\beta_0+\frac{\chi_q}{C}\right)\frac{d^2_{\alpha,\theta_1,\theta_2}(\rho_0)}{d_{2\alpha,\theta_1,\theta_2}(\rho_0)},$$

and with $\beta_q=\beta_0+\chi_q/C$, from (1.10), we get (6.3) and consistency follows.

The results presented in cases (i), (ii) and (iii) still hold true if we replace ρ_q by any consistent estimator of ρ_q, $\widehat{\rho}^{(q)}$, such that $(\widehat{\rho}^{(q)}-\rho_q)\ln(n/k)=o_p(1)$. $\square$

Remark 6.1. *The replacement of ρ by $\widehat{\rho}^{(q)}$ in the scale factor $2d_{2\alpha,\theta_1,\theta_2}(\rho_q)/\left(\alpha\eta_q d^2_{\alpha,\theta_1,\theta_2}(\rho)\right)$ places no problem due to a continuity argument, provided that $\widehat{\rho}^{(q)}$ is consistent for the estimation of ρ_q. However, the replacement of ρ_q by $\widehat{\rho}^{(q)}$ in $(k/n)^{\rho_q}$ requires that $(k/n)^{\rho_q}/(k/n)^{\widehat{\rho}^{(q)}} \underset{n\to\infty}{\overset{p}{\longrightarrow}} 1$ and hence the need to impose the condition $\left(\widehat{\rho}^{(q)}-\rho_q\right)\ln(n/k)=o_p(1)$.*

Appendix: The second and third-order frameworks for heavy-tailed models under a non-null shift

As mentioned above, if we induce any arbitrary shift, $s\in\mathbb{R}\backslash\{0\}$, in the unshifted model X_0, with quantile function $U_0(t)$, the transformed r.v., $X_s=X_0+s$, has an associated quantile function given by $U_s(t)=U_0(t)+s$. The second and third-order conditions in (1.4) and (1.5), respectively, can then be rewritten as

$$\lim_{t\to\infty}\frac{\ln U_s(tx)-\ln U_s(t)-\gamma\ln x}{A_s(t)}=\frac{x^{\rho_s}-1}{\rho_s} \tag{6.6}$$

and

$$\lim_{t\to\infty} \frac{\frac{\ln U_s(tx)-\ln U_s(t)-\gamma\ln x}{A_s(t)} - \frac{x^{\rho_s}-1}{\rho_s}}{B_s(t)} = \frac{x^{\rho_s+\rho'_s}-1}{\rho_s+\rho'_s}, \tag{6.7}$$

and hold for all $x > 0$, with $|A_s| \in RV_{\rho_s}$, $|B_s| \in RV_{\rho'_s}$, $\rho_s,\ \rho'_s < 0$. As a replacement of Lemma 3.1, if we assume that $U_0 \in RV_\gamma$ satisfies the second-order condition (1.4) with $\rho = \rho_0$ and $A = A_0$. Then $U_s(t) := U_0(t) + s$ is such that $U_s \in RV_\gamma$ and (6.6) holds with

$$A_s(t) := \begin{cases} -\gamma s/U_0(t), & \text{if} \quad \gamma+\rho_0 < 0 \ \wedge \ s \neq 0 \\ A_0(t) - \gamma s/U_0(t), & \text{if} \quad \gamma+\rho_0 = 0 \ \wedge \ s \neq 0 \\ A_0(t), & \text{if} \quad \gamma+\rho_0 > 0 \vee s = 0, \end{cases} \tag{6.8}$$

and (β_s, ρ_s) given in (1.7).

Consequently, the introduction of a shift in the model underlying the data can possibly change the shape second-order parameter $\rho = \rho_s$, in (1.4), which is indeed equal to $-\gamma$ whenever we induce a non-null shift in any unshifted model with $\gamma + \rho_0 < 0$, as, for instance, $X \equiv X_0 \frown$ Fréchet$(\gamma = 0.25)$, for which $\rho_0 = -1$. Then, and for $X_s = X_0 + s$, $s \neq 0$, the second-order parameter ρ, in (1.4), becomes $-\gamma$. In the sequel, and for a reasonably large set $\mathcal{H}$ of heavy-tailed models, $\mathcal{H} \subset \mathcal{D}_{\mathcal{M}}(G_{\gamma>0})$, we shall analyse the impact of a shift $s \neq 0$ not only on (β, ρ) and $A(\cdot)$, but, more generally, in the vector of unknown parameters $(\beta, \rho, \beta', \rho')$, with (β', ρ') the scale and shape third-order parameters, proceeding to a characterisation of $(\beta_s, \rho_s, \beta'_s, \rho'_s)$ and the functionals $U_s(t)$, $A_s(t)$ and $B_s(t)$, comparatively with the functionals $U_0(t)$, $A_0(t)$ and $B_0(t)$ corresponding to an unshifted model.

A subclasse of Hall-Welsh class of models

The so-called Hall-Welsh class of models was first introduced in Hall (1982), later used in Hall and Welsh (1985) with a restriction $E_1 \neq 0$, and it is now used under a third-order framework. We thus assume to be working in a class $\mathcal{H}$ of heavy-tailed models, such that

$$\overline{F}(x) \equiv \overline{F}_0(x) = (x/C)^{-1/\gamma}\Big\{1 + E_1\,(x/C)^{\rho_0/\gamma} + E_2\,(x/C)^{2\rho_0/\gamma} + o\big(x^{2\rho_0/\gamma}\big)\Big\},$$

as $x \to \infty$, where $\gamma > 0$, $\rho_0 < 0$, $C > 0$ and $E_1, E_2 \neq 0$. Equivalently, we can say that, as $t \to \infty$,

$$U(t) \equiv U_0(t) = Ct^\gamma\left\{1 + D_1 t^{\rho_0} + D_2 t^{2\rho_0} + o(t^{2\rho_0})\right\}, \tag{6.9}$$

where $D_1 = \gamma E_1$ and $D_2 = \gamma\left(E_2 + (2\rho_0 + \gamma - 1)/2E_1^2\right)$. Then, the third-order condition in (1.5) holds, with $\rho = \rho' = \rho_0$, $A_0(t) = \rho_0 D_1 t^{\rho_0}$ and $B_0(t) = (2D_2/D_1 - D_1)\, t^{\rho_0}$. For this class of models and choosing the parameterizations $A_0(t) := \gamma\beta_0 t^{\rho_0}$ and $B_0(t) := \beta_0' t^{\rho_0}$, we have $D_1 = \gamma\beta_0/\rho_0$ and $D_2 = D_1(D_1 + \beta_0')/2$.

Remark 6.2. *The* log-gamma *model* ($\rho = 0$ *in* (1.4)) *is out of the class of models in* (6.9)*. The unit* Pareto *model, with d.f.* $F(z) = 1 - z^{-1/\gamma}$*,* $z \geq 1$*, and quantile function* $U(t) = t^\gamma$*,* $t \geq 1$*, is also out of this class of models. Indeed, we get* $U(tx)/U(t) = x^\gamma$ *for all* $x \geq 1$*, i.e.* $A(t) \equiv 0$ *in* (1.4) *meaning that we may assume the fastest convergence attached to* $\rho = -\infty$*.*

If we induce a deterministic shift, $s \in \mathbb{R}\backslash\{0\}$, in the underlying model, the associated reciprocal quantile function, $U_s(t)$, is then given by,

$$U_s(t) = Ct^\gamma \left\{1 + D_1 t^{\rho_0} + D_2 t^{2\rho_0} + sC^{-1}t^{-\gamma} + o(t^{2\rho_0})\right\}, \quad \text{as} \quad t \to \infty.$$

The parameter $(\beta = \beta_s, \rho = \rho_s)$, in (1.6), is then the one given in (1.7). The function $A_s(t)$ depends thus on the relationship between the first-order parameter γ, and the second-order parameter ρ_0, just as provided in (6.8). The characterisation of $B_s(t)$ in (6.7) is slightly more complex, and it is presented, jointly with $A_s(t)$, in Table 1.

Table 1: Characterisation of second, third-order parameters and functionals A_s and B_s for a model F in the Hall-Welsh sub-class of models, in (6.9), additionally subject to a shift $s \neq 0$.

		$A_s(t) := \gamma\beta_s t^{\rho_s}$	$B_s(t) := \beta_s' t^{\rho_s'}$
$\gamma + \rho_0 < 0$	$\rho_0 < -2\gamma$	$-\frac{\gamma s}{C} t^{-\gamma}$	$-\frac{s}{C} t^{-\gamma}$
	$\rho_0 = -2\gamma$		$\left(\frac{2\gamma\beta_0 C}{\rho_0 s} - \frac{s}{C}\right) t^{-\gamma}$
	$\rho_0 > -2\gamma$		$-\frac{\beta_0 C}{s}\ t^{\gamma+\rho_0}$
$\gamma + \rho_0 = 0$		$\left(\gamma\beta_0 + \frac{\rho_0 s}{C}\right) t^{\rho_0}$	$\frac{\beta_0 C^2(\beta_0' - \beta_0) + 2(\beta_0 C - s)^2}{2C(\beta_0 C - s)} t^{\rho_0}$
$\gamma + \rho_0 > 0$	$\rho_0 < -\gamma/2$	$\gamma\beta_0 t^{\rho_0} \equiv A_0(t)$	$-\frac{s}{\beta_0 C} t^{-(\gamma+\rho_0)}$
	$\rho_0 = -\gamma/2$		$\left(\beta_0' + \frac{2\rho_0 s}{\gamma\beta_0 C}\right) t^{\rho_0}$
	$\rho_0 > -\gamma/2$		$\beta_0' t^{\rho_0} \equiv B_0(t)$

Remark 6.3. *The results presented in* Table 1 *enable us to fully characterise any model in the aforementioned sub-class of* Hall-Welsh*'s class*:

- *For the* Burr(γ, ρ) *model with d.f.* $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$ $(x > 0,\ \gamma > 0,\ \rho \equiv \rho_0 < 0)$ *we have* $C = 1$, $D_1 = \gamma/\rho_0$ *and* $D_2 = D_1(1 + D_1)/2$.
- *For the* Fréchet(γ) *model with d.f.* $F(x) = \exp\left(-x^{-1/\gamma}\right)$ $(x > 0,\ \gamma > 0)$, *we have* $\rho_0 = -1$, $C = 1$, $D_1 = -\gamma/2$ *and* $D_2 = D_1\left(5/6 + D_1\right)/2$.
- *For the* generalised Pareto (GP)$(\gamma > 0)$ *model with d.f.* $F(x) = 1 - (1 + \gamma x)^{-1/\gamma}$, $(x > 0,\ \gamma > 0)$, *we have* $\rho_0 = -\gamma$, $C = 1/\gamma$, $D_1 = -1$ *and* $D_2 = 0$.
- *The* Student*'s-*t_ν $(\nu > 0)$ *distribution is*

$$F(x) = F(x|\nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}} \int_{-\infty}^{x} \left(1 + \frac{z^2}{\nu}\right)^{-(\nu+1)/2} dz,\ x \in \mathbb{R},\ \nu > 0,$$

with $\gamma = 1/\nu$ *and* $\rho_0 = -2/\nu = -2\gamma$. *In this case we have,* $C = \sqrt{\nu}/c_\nu$, *where* $c_\nu = (\nu \mathcal{B}(\nu/2, 1/2))^{1/\nu}$, *with* $\mathcal{B}$ *the complete Beta function, and* $D_1 = -c_\nu^2(\nu+1)/(2(\nu+2))$. *When* $\nu = 1$ *we get the so called* Cauchy *d.f.,* $F(x) = 1/2 - (\arctan(x))/\pi,\ x \in \mathbb{R}$, *with* $\gamma = 1$ *and* $\rho_0 = -2$. *For the* Cauchy *distribution we have* $C = 1/\pi$ *and* $D_1 = -\pi^2/3$.

Remark 6.4. *Just as mentioned in* Remark 3.1, *note that we can use in* (6.6), (6.7) *and* (6.8) *the subscript* q *instead of the subscript* s, *whenever we think on such a shift as* $s = -\chi_q$.

Acknowledgement Research partially supported by National Funds through **FCT** — Fundação para a Ciência e a Tecnologia, project PEst-OE/MAT/UI0006/2011, PTDC/FEDER, EXTREMA project and grant SFRH/BPD/77319/2011.

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., Caeiro, F. and Gomes, M.I. (2012). An overview and open research topics in the field of statistics of univariate extremes. *Revstat* **10**:1, 1-31.

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

[4] Dekkers, A., Einmahl, J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Annals of Statistics* **17**, 1833–1855.

[5] 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.

[6] Geluk, J. and de Haan, L. (1987). *Regular Variation, Extensions and Tauberian Theorems.* CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, The Netherlands.

[7] Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics* **44**, 423–453.

[8] Gomes, M.I. (2003). Stochastic processes in telecommunication traffic. In Fernandes, C. *et al.* (eds.), *Mathematical Techniques and Problems in Telecommunications*, CIM edition, 7–32.

[9] Gomes, M.I. and Martins, M.J. (2002). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *J. Statist. Planning and Inference* **93**, 161–180.

[10] Gomes, M.I., de Haan, L. and Peng, L. (2002). Semi-parametric estimation of the second order parameter — asymptotic and finite sample behaviour. *Extremes* **5**:4, 387–414.

[11] Gomes, M.I., Canto e Castro, L., Fraga Alves, M.I. and Pestana, D. (2008a). Statistics of extremes for iid data and breakthroughs in the estimation of the extreme value ndex: Laurens de Haan leading contributions. *Extremes* **11**:1, 3–34.

[12] Gomes, M.I., Fraga Alves, M.I. and Araújo Santos, P. (2008b). PORT Hill and moment estimators for heavy-tailed models. *Commun. in Statist. – Simul. and Comput.* **37**, 1281–1306.

[13] Gomes, M.I., Henriques-Rodrigues, L. and Miranda, C. (2011a). Reduced-bias location-invariant extreme value index estimation: a simulation study. *Commun. in Statist. – Simul. and Comput.* **40**:3, 424–447.

[14] Gomes, M.I., Henriques-Rodrigues, L., Fraga Alves, M.I. and Manjunath, B.G. (2011b). Adaptive PORT-MVRB estimation: an empirical comparison of two heuristic algorithms. *J. Statist. Comput. Simul.*, in press. DOI:10.1080/00949655.2011.652113

[15] Haan, L. de (1984). Slow variation and characterization of domains of attraction. In Tiago de Oliveira, ed., *Statistical Extremes and Applications*, 31-48, D. Reidel, Dordrecht, Holland.

[16] Hall, P. (1982). On estimating the endpoint of a distribution. *Ann. Statist.* **10**, 556–568.

[17] Hall, P. and Welsh, A.W. (1985). Adaptive estimates of parameters of regular variation. *Annals of Statistics* **13**, 331–341.

[18] Henriques-Rodrigues, L. and Gomes, M.I. (2009). High quantile estimation and the PORT methodology. *Revstat* **7**:3, 245–264.

[19] Henriques-Rodrigues, L. and Gomes, M.I. (2012). A note on the PORT methodology in the estimation of a shape second-order parameter. In Oliveira, P., Temido, M.G., Henriques, C. and Vichi, M. (eds.) *Studies in Theoretical and Applied Statistics: Subseries B: Recent Developments in Modeling and Applications in Statistics* (SPE2010). Springer, 127–137.

[20] Henriques-Rodrigues, L., Gomes, M.I., Fraga Alves, M.I. and Neves, C. (2012). PORT-estimation of a Shape Second-order Parameter. Notas e Comunicações CEAUL 09/2012.

[21] Hüsler, J. (2009). Extreme value analysis in biometrics. *Biometrical Journal* **51**:2, 252–272.

[22] Reiss, R.-D. and Thomas, M. (2001; 2007). *Statistical Analysis of Extreme Values, with Application to Insurance, Finance, Hydrology and Other Fields*, 2nd edition, 3rd edition, Birkhäuser Verlag.

[23] Resnick, S.I. (1997). Heavy tail modelling and teletraffic data. *Annals of Statistics* **25**, 1805–1869.

[24] van der Vaart, A.W. (1998). *Asymptotic Statistics.* Cambridge University Press.