

Asymptotic Comparison of Alternative Estimators of a Shape Second-Order Parameter*

Frederico Caeiro

Universidade Nova de Lisboa, DM and CMA, e-mail: fac@fct.unl.pt

M. Ivette Gomes

Universidade de Lisboa, FCUL, DEIO and CEAUL, e-mail: ivette.gomes@fc.ul.pt

January 2, 2013

Abstract

Under a third-order framework, and for heavy right tails, or equivalently a positive extreme value index, we proceed to an asymptotic comparison of three alternative estimators of the most common shape second-order parameter associated to the type of models under consideration.

Keywords. Asymptotic behaviour; extreme value theory; semi-parametric estimation; shape second-order parameter; third-order frameworks.

1 Introduction and scope of the paper

In *statistics of extremes*, a model F is said to have a *heavy right tail* whenever the right tail-function $\bar{F} := 1 - F$ is a *regularly varying function* at infinity, with a negative index of regular variation denoted by $-1/\gamma$, $\gamma > 0$, or equivalently, and using the notation $F^{\leftarrow}(x) := \inf\{y : F(y) \geq x\}$ for the generalised inverse function of F , the

*Research partially supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, projects PEst-OE/MAT/UI0006/2011 (CEAUL), PEst-OE/MAT/UI0297/2011 (CMA/UNL) and PTDC/FEDER, EXTREMA.

(reciprocal) quantile function $U(t) := F^{-1}(1 - 1/t)$, $t \geq 1$, is of regular variation with an index γ . Recall that this means that for all $x > 0$,

$$\lim_{t \rightarrow \infty} \bar{F}(tx)/\bar{F}(t) = x^{-1/\gamma} \iff \lim_{t \rightarrow \infty} U(tx)/U(t) = x^\gamma \quad (1)$$

(see Bingham *et al.*, 1987, for details on regular variation).

The second-order parameter $\rho (\leq 0)$ measures the rate of convergence in the first-order condition in (1), and it is the non-positive parameter in the limiting relation,

$$\lim_{t \rightarrow \infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \begin{cases} \frac{x^\rho - 1}{\rho} & \text{if } \rho < 0 \\ \ln x & \text{if } \rho = 0 \end{cases} =: \psi_\rho(x), \quad (2)$$

which we assume to hold for all $x > 0$, and where $|A|$ must be of regular variation with an index ρ (Geluk and de Haan, 1987). Under a framework of a heavy right tail-function $\bar{F} = 1 - F$, with F the model underlying the available data, we proceed to an asymptotic comparison of three alternative estimators of the shape second-order parameter ρ , in (2), the implicit estimator in Feuerverger and Hall (1999), one of the classes of estimators in Fraga Alves *et al.* (2003), also included and found competitive in the more recent papers by Goegebeur *et al.* (2008; 2010), Ciuperca and Mercadier (2010) and Caeiro and Gomes (2012a,b), and the estimator introduced and studied in Caeiro and Gomes (2012a), all on the ρ -estimation.

In order to get full information on the asymptotic non-degenerate behaviour of ρ -estimators, it is often necessary to further assume a third-order condition, ruling the rate of convergence in (2). We shall assume that for all $x > 0$,

$$\lim_{t \rightarrow \infty} \frac{\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} - \psi_\rho(x)}{B(t)} = \begin{cases} \frac{x^{\rho+\rho'} - 1}{\rho+\rho'} & \text{if } \rho \vee \rho' < 0 \\ \ln x & \text{if } \rho = \rho' = 0, \end{cases} \quad (3)$$

where $|B(t)|$ must then be of regular variation with an index ρ' . We are then confronted with this third-order parameter $\rho' \leq 0$. In this article, we shall more restrictively assume that we are working with a Paretian type class of models, with a right tail-function given by

$$\bar{F}(x) = 1 - F(x) = Cx^{-1/\gamma} \{1 + D_1 x^{\rho/\gamma} + D_2 x^{2\rho/\gamma} + o(x^{2\rho/\gamma})\} \quad (4)$$

as $x \rightarrow \infty$, with $C > 0$, $D_1, D_2 \neq 0$, $\rho < 0$. Note that to assume such a kind of right tail-function is equivalent to say that (3) holds with $\rho = \rho' < 0$ and that we can

choose the same parameterisation as in Caeiro and Gomes (2011),

$$A(t) = \alpha t^\rho =: \gamma \beta t^\rho, \quad B(t) = \beta' t^\rho = \frac{\beta' A(t)}{\beta \gamma} =: \frac{\xi A(t)}{\gamma}, \quad \xi = \frac{\beta'}{\beta}, \quad (5)$$

with $\beta \neq 0$ and $\beta' \neq 0$ “scale” second and third-order parameters, respectively.

In Section 2 of this paper, we shall introduce the ρ -estimators under analysis. Section 3 is dedicated to the description of the asymptotic behaviour of such estimators under the third-order framework in (4). Finally, in Section 4, we deal with an asymptotic comparison of the ρ -estimators under consideration for a fixed k , the number of top order statistics involved in the estimation, and at optimal levels, i.e., levels k_1 where the asymptotic mean square error (MSE) of the ρ -estimators is minimised.

2 The ρ -estimators under study

Given a random sample, (X_1, \dots, X_n) , and with the notation $(X_{1:n} \leq \dots \leq X_{n:n})$ for the sample of the associated ascending order statistics (o.s.'s), one of the classes of ρ -estimators under analysis is the simplest class introduced and studied both asymptotically and for finite samples in Fraga Alves *et al.* (2003). As shown in Caeiro and Gomes (2006), such a class of estimators can be parameterised in a tuning or control parameter τ not necessarily non-negative, but real, and it is given by

$$\hat{\rho}_n^{\text{FAGH}}(k) \equiv \hat{\rho}_n^{\text{FAGH}(\tau)}(k) := - \left| \frac{3(T_n^{(\tau)}(k) - 1)}{T_n^{(\tau)}(k) - 3} \right|, \quad (6)$$

with

$$T_n^{(\tau)}(k) := \frac{(M_n^{(1)}(k))^\tau - (M_n^{(2)}(k)/2)^{\tau/2}}{(M_n^{(2)}(k)/2)^{\tau/2} - (M_n^{(3)}(k)/6)^{\tau/3}},$$

with the notation $a^\tau = \ln a$ whenever $\tau = 0$ and where, with V_{ik} the scaled log-excesses, i.e.,

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \leq i \leq k,$$

we have

$$M_n^{(j)}(k) := \frac{1}{k} \sum_{i=1}^k V_{ik}^j, \quad j \geq 1.$$

Remark 1. Note that, in the notation used in Fraga Alves *et al.* (2003), $\hat{\rho}_n^{\text{FAGH}(\tau)}(k) = \hat{\rho}_{n|T}^{(\alpha, \theta_1, \theta_2, \tau)}(k)$, for the tuning parameters $(\alpha, \theta_1, \theta_2) = (1, 2, 3)$.

More recently, Caeiro and Gomes (2012a) considered consistent estimators of $\gamma > 0$ defined by adequate linear combinations of the scaled log-spacings

$$U_i := i\{\ln X_{n-i+1:n} - \ln X_{n-i:n}\}, \quad 1 \leq i \leq k < n, \quad (7)$$

given by

$$N_{n,k}^{(\alpha)} := \frac{\alpha}{k} \sum_{i=1}^k \left(\frac{i}{k}\right)^{\alpha-1} U_i, \quad \alpha \geq 1,$$

and proposed a new class of ρ -estimators with the functional expression,

$$\hat{\rho}_n^{CG}(k) = \hat{\rho}_n^{CG(\tau)}(k) := \min \left\{ 0, 1 + \frac{1}{1 - R_{n,k}^{(\tau)}} \right\}, \quad \tau \in \mathbb{R}, \quad (8)$$

where

$$R_{n,k}^{(\tau)} = \frac{\left(N_{n,k}^{(1)}\right)^\tau - \left(N_{n,k}^{(3/2)}\right)^\tau}{\left(N_{n,k}^{(3/2)}\right)^\tau - \left(N_{n,k}^{(2)}\right)^\tau}, \quad \tau \in \mathbb{R},$$

again with the notation $a^\tau = \ln a$ whenever $\tau = 0$.

Remark 2. *As mentioned in Caeiro and Gomes (2012a), note that we could also have worked with*

$$R_{n,k}^{(\tau, \alpha_1, \alpha_2, \alpha_3)} = \frac{\left(N_{n,k}^{(\alpha_1)}\right)^\tau - \left(N_{n,k}^{(\alpha_2)}\right)^\tau}{\left(N_{n,k}^{(\alpha_2)}\right)^\tau - \left(N_{n,k}^{(\alpha_3)}\right)^\tau}, \quad \tau \in \mathbb{R},$$

$\alpha_i \neq \alpha_j$, $1 \leq i < j \leq 3$ and $\min_{1 \leq i \leq 3}(\alpha_i) \geq 1$. But then we had to deal with the choice of the values of additional tuning parameters.

The third class of ρ -estimators is based on the fact that in Hall-Welsh's class of models (Hall, 1982; Hall and Welsh, 1985), with a right tail function

$$\bar{F}(x) = Cx^{-1/\gamma} \{1 + Dx^{\rho/\gamma} + o(x^\rho)\},$$

as $x \rightarrow \infty$, $C > 0$, $D \neq 0$, $\rho < 0$, the log-spacings U_i , $1 \leq i \leq k$, in (7), are approximately exponential with mean value

$$\mu_i = \gamma e^{\beta(i/n)^{-\rho}}, \quad 1 \leq i \leq k.$$

Feuerverger and Hall (1999) considered the joint maximization, in γ , β and ρ , of the log-likelihood of the scaled log-spacings, U_i , $1 \leq i \leq k$, in (7), given by

$$\ln L(\gamma, \beta, \rho; U_i, 1 \leq i \leq k) = -k \ln \gamma - \beta \sum_{i=1}^k \left(\frac{i}{n}\right)^{-\rho} - \frac{1}{\gamma} \sum_{i=1}^k e^{-\beta(i/n)^{-\rho}} U_i.$$

Such a maximization leads to implicit estimators $\hat{\beta}$ and $\hat{\rho}$, such that

$$(\hat{\beta}, \hat{\rho}) := \arg \min_{(\beta, \rho)} \left\{ \ln \left(\frac{1}{k} \sum_{i=1}^k e^{-\beta(i/n)^{-\rho}} U_i \right) + \beta \left(\frac{1}{k} \sum_{i=1}^k \left(\frac{i}{n} \right)^{-\rho} \right) \right\}. \quad (9)$$

Such a class of implicit ρ -estimators will be denoted $\hat{\rho}_n^{\text{FH}}(k)$.

3 Asymptotic behaviour of the ρ -estimators

Let us consider the usual notation $\text{Normal}(\mu, \sigma^2)$ for a random variable (r.v.) with mean value μ and variance σ^2 . On the basis of the research performed in Fraga Alves *et al.* (2003), for the ρ -estimators $\hat{\rho}_n^{\text{FAGH}}(k)$, in (6), in Caeiro and Gomes (2012a), for the ρ -estimators $\hat{\rho}_n^{\text{CG}}(k)$, in (8), and in Caeiro and Gomes (2011), regarding $\hat{\rho}_n^{\text{FH}}(k)$, the implicit ρ -estimator in Feuerverger and Hall (1999) obtained through (9), but with some further computations related to the asymptotic bias of the estimators under analysis, performed in Caeiro *et al.* (2009) and in this article, we now state the following theorem.

Theorem 1. *Under the validity of the second-order condition, in (2), with $\rho < 0$, and for intermediate values of k , i.e., k -values such that*

$$k = k_n \rightarrow \infty, \quad k/n \rightarrow 0, \quad \text{as } n \rightarrow \infty,$$

if we further assume that $\sqrt{k} A(n/k) \rightarrow \infty$, as $n \rightarrow \infty$, any of the classes of estimators, $\hat{\rho}_n^{\text{FAGH}}(k) \equiv \hat{\rho}_n^{\text{FAGH}(\tau)}(k)$, $\hat{\rho}_n^{\text{CG}}(k) \equiv \hat{\rho}_n^{\text{CG}(\tau)}(k)$, $\tau \in \mathbb{R}$ and $\hat{\rho}_n^{\text{FH}}(k)$, given in (6), (8) and (9), respectively, are consistent for the estimation of ρ .

If we further assume the validity of the third-order condition in (3), A and B given in (5), and with U_k^{FH} , U_k^{FAGH} and U_k^{CG} asymptotically standard normal r.v.'s, we can guarantee that

$$\hat{\rho}_n^\bullet(k) - \rho \stackrel{d}{=} \frac{\sigma_\bullet U_k^\bullet}{\sqrt{k} A(n/k)} + b_\bullet A(n/k)(1 + o_p(1)), \quad (10)$$

where

$$\sigma_{\text{FH}} = \gamma (1 - \rho)(1 - 2\rho)\sqrt{1 - 2\rho/|\rho|}, \quad (11)$$

$$\sigma_{\text{FAGH}} = \gamma(1 - \rho)^3\sqrt{2\rho^2 - 2\rho + 1/|\rho|}, \quad (12)$$

$$\sigma_{\text{CG}} = \gamma(1 - \rho)(2 - \rho)(3 - 2\rho)\sqrt{4\rho^2 - 4\rho + 7}/(\sqrt{120}|\rho|), \quad (13)$$

$$b_{\text{FH}} = \frac{(2\xi - 1)(1 - \rho)^2}{2\gamma(1 - 2\rho)(1 - 3\rho)^2}, \quad (14)$$

$$b_{\text{FAGH}} = \frac{\tau\rho(3 - \rho)(3 - 2\rho)(1 - 2\rho)^2 - 6\rho^2(4\rho^3 - 16\rho^2 + 20\rho - 7)}{12\gamma(1 - \rho)^2(1 - 2\rho)^2} + \frac{2\rho\xi(1 - \rho)^3}{\gamma(1 - 2\rho)^3}, \quad (15)$$

$$b_{\text{CG}} = \frac{(\tau - 1)\rho}{2\gamma} + \frac{\rho\xi(2 - \rho)(3 - 2\rho)}{\gamma(1 - 2\rho)(3 - 4\rho)}. \quad (16)$$

Consequently, for levels k such that $\sqrt{k} A^2(n/k) \rightarrow \lambda_A$, finite, there exist real constants b_{FH} , b_{FAGH} , b_{CG} , defined in (14), (15) and (16), respectively, and generally denoted b_{\bullet} , and positive real constants σ_{FH} , σ_{FAGH} , σ_{CG} , defined in (11), (12), (13), respectively, generally denoted σ_{\bullet} , such that

$$\sqrt{k} A(n/k)(\hat{\rho}_n^{\bullet}(k) - \rho) \xrightarrow[n \rightarrow \infty]{\text{d}} \text{Normal}(\lambda_A b_{\bullet}, \sigma_{\bullet}^2). \quad (17)$$

Remark 3. The asymptotic variance σ_{FH} , in (11), was computed in Caeiro and Gomes (2011). The asymptotic variance σ_{FAGH} , in (12), was explicitly computed in Fraga Alves et al. (2003) and the asymptotic variance σ_{CG} , in (13), was explicitly computed in Caeiro and Gomes (2012a). For models such that the third-order condition in (3) holds, with A and B given in (5), the asymptotic bias b_{FH} , in (14), was also derived in Caeiro and Gomes (2011). The asymptotic bias b_{FAGH} , in (15), was given in Fraga Alves et al. (2003), and explicitly computed in Caeiro et al. (2009), for models with ρ possibly different from ρ' . Finally, the asymptotic bias b_{CG} , in (16), was explicitly computed in Caeiro and Gomes(2012a), also for models with ρ possibly different from ρ' , with (ρ, ρ') the second-order parameters in (3).

4 Asymptotic comparison of the ρ -estimators

We shall now proceed to an asymptotic comparison of the estimators under study, generally denoted $\hat{\rho}_n^{\bullet}(k)$, not only for any general k , but also at optimal levels, i.e., at levels $k = k_1^{\bullet}$ where the asymptotic mean square error of $\hat{\rho}_n^{\bullet}(k)$ is minimum.

4.1 Asymptotic comparison at a level k

Regarding the asymptotic standard deviations σ_{FH} and σ_{FAGH} , in (11) and (12), respectively, note that since $\rho < 0$, we have $\sigma_{\text{FH}} < \sigma_{\text{FAGH}}$, $\forall \rho < 0$, with

$$\frac{\sigma_{\text{FH}}}{\sigma_{\text{FAGH}}} = \frac{(1-2\rho)\sqrt{1-2\rho}}{(1-\rho)\sqrt{2\rho^2-2\rho+1}}$$

approaching zero, as $|\rho| \rightarrow \infty$, and equal to 1 for $\rho = 0$. Regarding the asymptotic standard deviations σ_{FH} and σ_{CG} , in (11) and (13), respectively, we have

$$\frac{\sigma_{\text{FH}}}{\sigma_{\text{CG}}} = \frac{(1-2\rho)\sqrt{120(1-2\rho)}}{(2-\rho)(3-2\rho)\sqrt{4\rho^2-4\rho+7}},$$

smaller than one for all $\rho \leq 0$, approaching zero, as $|\rho| \rightarrow \infty$, with a maximum close to one at $\rho = -0.725$ and approaching the value $\sqrt{210}/21$ as ρ approaches zero. In Figure 1 we illustrate such a behaviour.

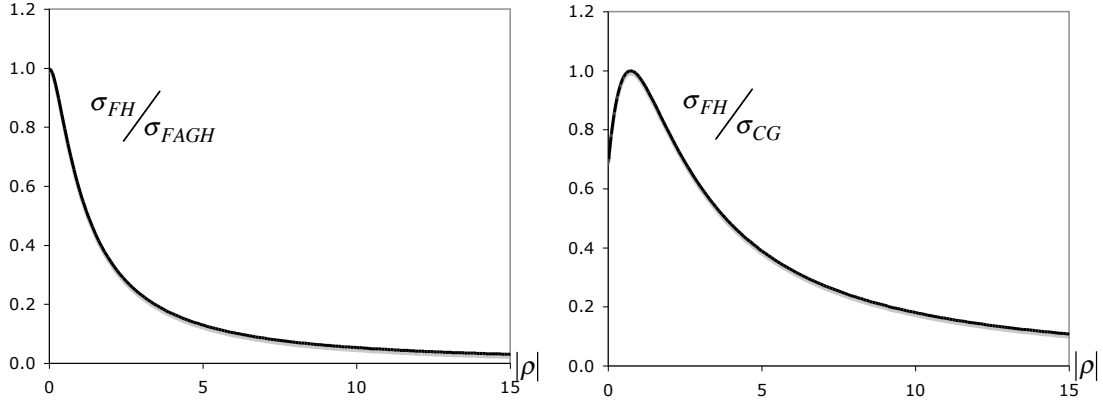


Figure 1: Values of $\sigma_{\text{FH}}/\sigma_{\text{FAGH}}$ (left) and $\sigma_{\text{FH}}/\sigma_{\text{CG}}$ (right), as a function of $|\rho|$

The other way round, if we think on the bias b_{FH} and b_{FAGH} , given in (14) and (15), respectively, we get the quotient

$$\frac{b_{\text{FAGH}}}{b_{\text{FH}}} = \frac{(1-2\rho)(1-3\rho)^2}{3(2\xi-1)} \left(\frac{\tau\rho(3-\rho)(3-2\rho)}{2(1-\rho)^4} - \frac{3\rho^2(4\rho^3-16\rho^2+20\rho-7)}{(1-\rho)^4(1-2\rho)^2} + \frac{12\rho\xi(1-\rho)}{(1-2\rho)^3} \right),$$

and due to the fact that $\rho < 0$, if $\xi \neq 1/2$, there is always a value of τ that leads us to $b_{\text{FAGH}}/b_{\text{FH}} = 0$. Such a value, given by

$$\tau_0^{\text{FAGH}} \equiv \tau_0^{\text{FAGH}}(\xi, \rho) = \frac{6\rho(4\rho^3-16\rho^2+20\rho-7)(1-2\rho)-24\xi(1-\rho)^5}{(1-2\rho)^3(3-\rho)(3-2\rho)}, \quad (18)$$

approaches $3(2 - \xi)/2$, as $\rho \rightarrow -\infty$, and $-8\xi/3$, as $\rho \rightarrow 0$. If we think on the bias b_{FH} and b_{CG} , given in (14) and (16), respectively, we get the quotient

$$\frac{b_{\text{CG}}}{b_{\text{FH}}} = \left(\frac{(\tau - 1)\rho}{2} + \frac{\rho\xi(2 - \rho)(3 - 2\rho)}{(1 - 2\rho)(3 - 4\rho)} \right) \frac{2(1 - 2\rho)(1 - 3\rho)^2}{(2\xi - 1)(1 - \rho)^2}$$

and due to the fact that $\rho < 0$, if $\xi \neq 1/2$, there is also always a value of τ that leads us to $b_{\text{CG}}/b_{\text{FH}} = 0$, given by

$$\tau_0^{\text{CG}} \equiv \tau_0^{\text{CG}}(\xi, \rho) = 1 - \frac{2\xi(2 - \rho)(3 - 2\rho)}{(1 - 2\rho)(3 - 4\rho)}, \quad (19)$$

a value that approaches $-\xi/2$, as $\rho \rightarrow -\infty$, and $1 - 4\xi$, as ρ approaches zero. These facts were already noticed in Caeiro and Gomes (2012b), where bias reduction is considered.

One of the most typical values of ξ is $\xi = 1$, associated to models like the generalised Pareto (GP) model, with d.f. $F(x) = 1 - (1 + \gamma x)^{-1/\gamma}$, $x \geq 0$, $\gamma > 0$, and the Burr model, with d.f. $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x \geq 0$, $\gamma > 0$. For a Fréchet parent, another typical heavy-tailed model, we get $\xi = 10/6$. Finally, for the Student's t_ν -model with ν degrees of freedom, with a probability density function

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

we get $\gamma = 1/\nu$ and $\rho' = \rho = -2/\nu$. For an explicit expression of β and β' as a function of ν , see Caeiro and Gomes (2008). We have

$$\xi = \beta'/\beta = (\nu^2 + 4\nu + 2)/((\nu + 1)(\nu + 4)) \in (0.5, 1)$$

(see Caeiro and Gomes, 2011). In Figure 2 we represent graphically $\tau_0^{\text{FAGH}}(\xi, \rho)$, in (18), for a few values of ξ . Similarly, in Figure 3 we represent graphically $\tau_0^{\text{CG}}(\xi, \rho)$, in (19), for a few values of ξ .

The same values $\tau_0^{\text{FAGH}}(\xi, \rho)$ and $\tau_0^{\text{CG}}(\xi, \rho)$ are next presented at Figures 4 and 5, respectively, in the (ξ, ρ) -plane.

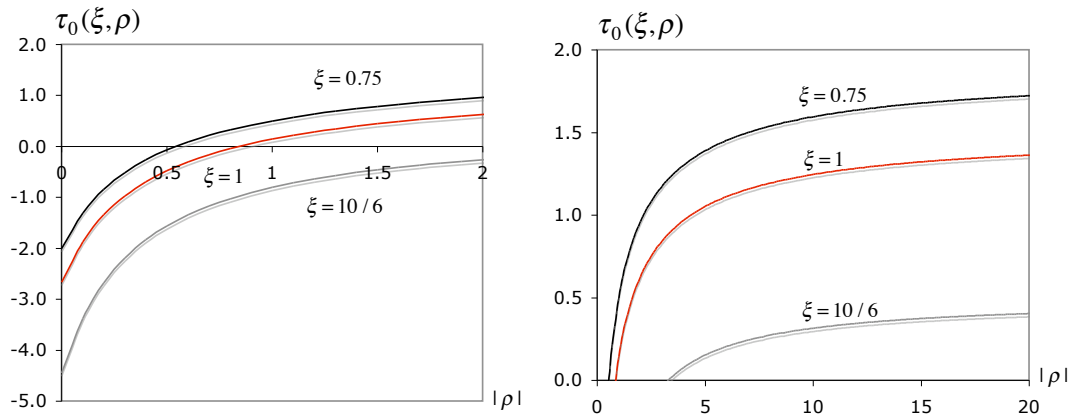


Figure 2: Values of $\tau_0(\xi, \rho) = \tau_0^{FAGH}(\xi, \rho)$, as a function of $|\rho|$ for $\xi = 0.75, 1$ and $10/6$, in two different scales

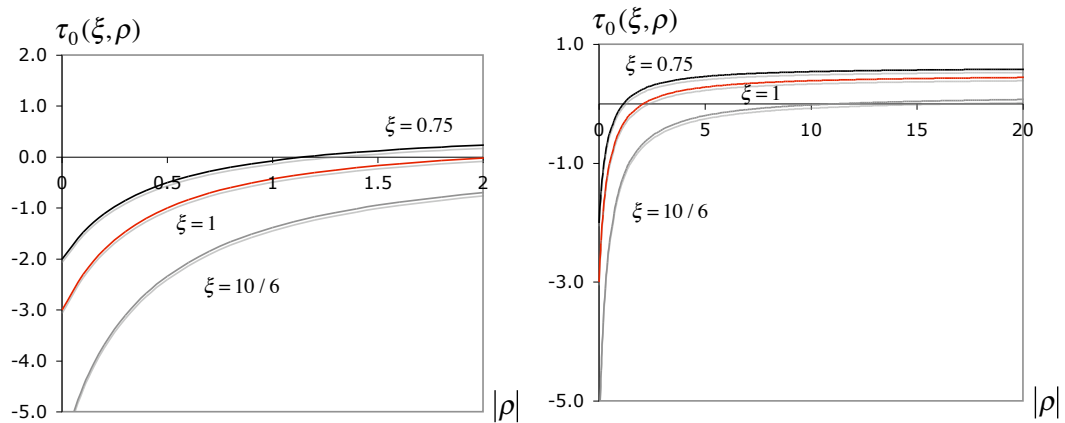


Figure 3: Values of $\tau_0(\xi, \rho) = \tau_0^{CG}(\xi, \rho)$, as a function of $|\rho|$ for $\xi = 0.75, 1$ and $10/6$, in two different scales

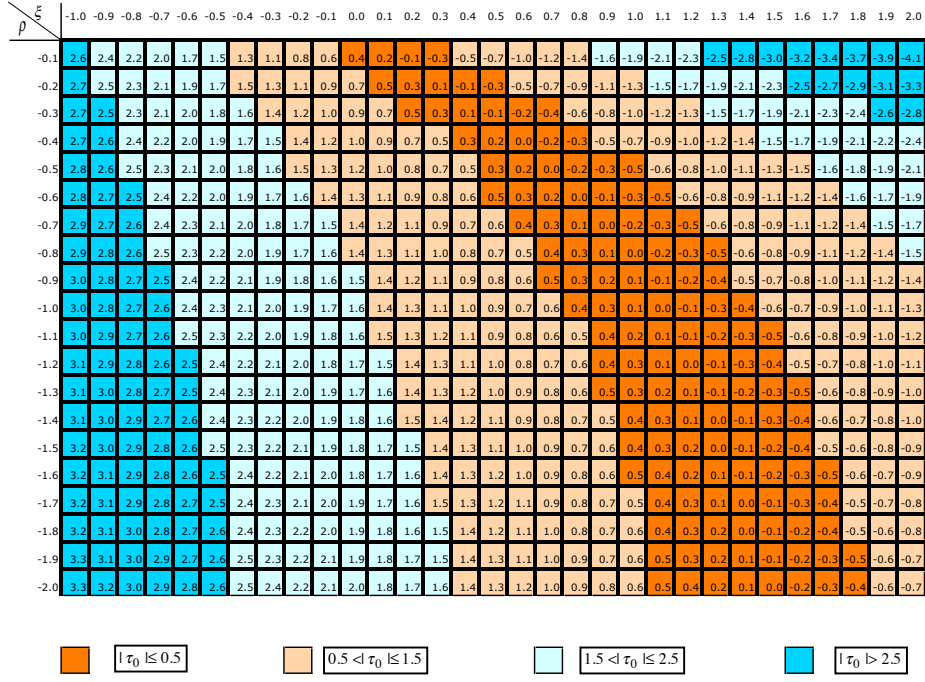


Figure 4: Values $\tau_0 = \tau_0^{FAGH}(\xi, \rho)$ such that $b_{FAGH} = 0$

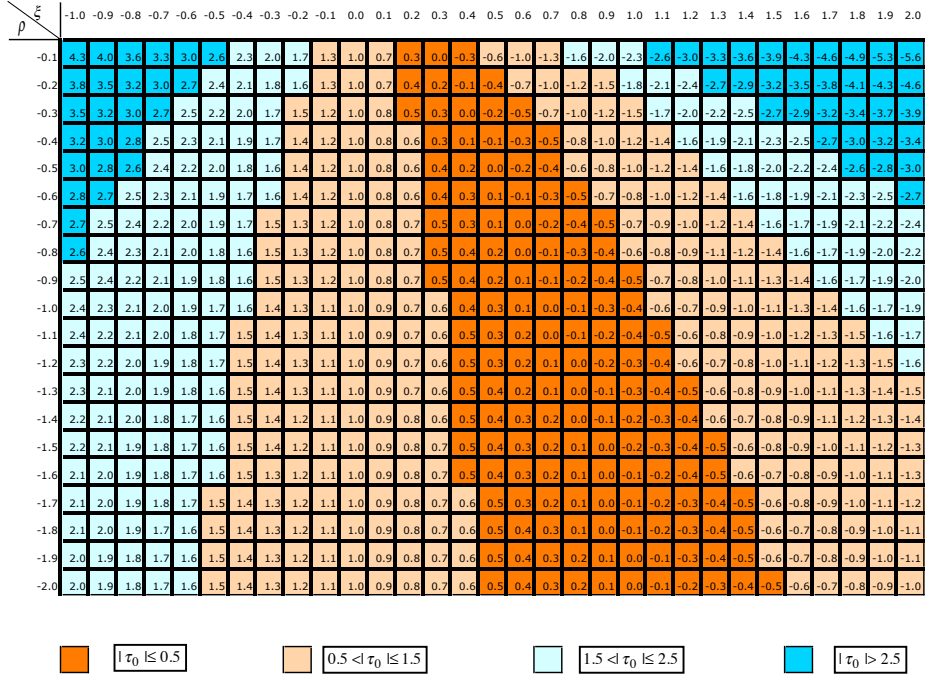


Figure 5: Values $\tau_0 = \tau_0^{CG}(\xi, \rho)$ such that $b_{CG} = 0$

4.2 Asymptotic comparison at optimal levels

We shall next proceed to the comparison of the estimators under study at their optimal levels. This is done in a way similar to the one used in de Haan and Peng (1998), Gomes and Martins (2001), Gomes *et al.* (2005, 2007b), Gomes and Neves (2008) and Gomes and Henriques-Rodrigues (2010) for the classical EVI-estimators, in Gomes *et al.* (2007a) for MVRB maximum likelihood EVI-estimators, and in Caeiro and Gomes (2011) for a larger set of MVRB EVI-estimators.

On the basis of the asymptotic distribution of $\hat{\rho}_n^\bullet(k)$, in (10), its asymptotic MSE (AMSE) is given by

$$\text{AMSE}_\bullet(k) = \frac{\sigma_\bullet^2}{kA^2(n/k)} + b_\bullet^2 A^2(n/k).$$

With the parameterisation in (5), we can write

$$\begin{aligned} \text{AMSE}_\bullet(k) &= \frac{\sigma_\bullet^2}{k\gamma^2\beta^2(n/k)^{2\rho}} + b_\bullet^2 \gamma^2\beta^2(n/k)^{2\rho} \\ &= \frac{1}{\gamma^2\beta^2} (\sigma_\bullet^2 n^{-2\rho} k^{2\rho-1} + b_\bullet^2 \gamma^4 \beta^4 n^{2\rho} k^{-2\rho}). \end{aligned}$$

The value $k_1^\bullet := \arg \min_k \text{AMSE}_\bullet(k)$ is solution in k of the equation,

$$-(1-2\rho)\sigma_\bullet^2 n^{-2\rho} k^{2\rho-2} + (-2\rho)b_\bullet^2 \gamma^4 \beta^4 n^{2\rho} k^{-2\rho-1} = 0.$$

We thus get

$$\frac{(1-2\rho)\sigma_\bullet^2}{(-2\rho)b_\bullet^2 \gamma^4 \beta^4} n^{-4\rho} = (k_1^\bullet)^{1-4\rho}$$

i.e. $\text{AMSE}_\bullet(k)$ is minimised at a level

$$k_1^\bullet := \arg \inf_k \text{AMSE}_\bullet(k) = \left(\frac{\sigma_\bullet^2(1-2\rho)}{b_\bullet^2 \gamma^4 \beta^4 (-2\rho)} \right)^{1/(1-4\rho)} n^{-4\rho/(1-4\rho)}. \quad (20)$$

Let us use the notation $\hat{\rho}_{n_0}^\bullet := \hat{\rho}_n^\bullet(k_1^\bullet)$, with k_1^\bullet given in (20).

On the basis of (20), we can write

$$\left(\frac{n}{k_1^\bullet} \right)^{2\rho} = \left(\frac{\sigma_\bullet^2(1-2\rho)}{b_\bullet^2 \gamma^4 \beta^4 (-2\rho)} \right)^{-2\rho/(1-4\rho)} n^{2\rho/(1-4\rho)}.$$

Let us use the notation

$$\varphi(\rho) := \left(\frac{1-2\rho}{-2\rho} \right)^{-(1-2\rho)/(1-4\rho)} + \left(\frac{1-2\rho}{-2\rho} \right)^{-2\rho/(1-4\rho)}.$$

We are now interested in the computation of

$$\begin{aligned} \text{AMSE}_{\bullet}(k_1^{\bullet}) &= \frac{\sigma_{\bullet}^2}{k_1 \gamma^2 \beta^2 (n/k_1^{\bullet})^{2\rho}} + b_{\bullet}^2 \gamma^2 \beta^2 (n/k_1^{\bullet})^{2\rho} \\ &= (\gamma^2 \beta^2)^{1/(1-4\rho)} (\sigma_{\bullet}^2)^{-2\rho/(1-4\rho)} (b_{\bullet}^2)^{(1-2\rho)/(1-4\rho)} \varphi(\rho) n^{2\rho/(1-4\rho)}. \end{aligned}$$

Consequently, and with the notation $\text{AMSE}(\hat{\rho}_{n0}^{\bullet}) := \text{AMSE}_{\bullet}(k_1^{\bullet})$, we can guarantee that whenever $b_{\bullet} \neq 0$, there exists a function $\psi(n) = \psi(n, \gamma, \beta, \rho)$, such that

$$\lim_{n \rightarrow \infty} \psi(n) \text{AMSE}(\hat{\rho}_{n0}^{\bullet}) = (\sigma_{\bullet}^2)^{-\frac{2\rho}{1-4\rho}} (b_{\bullet}^2)^{\frac{1-2\rho}{1-4\rho}} =: \text{LMSE}(\hat{\rho}_{n0}^{\bullet}).$$

It is then sensible to consider the following definition of *asymptotic root efficiency* (AREFF).

Definition 1. *Given two biased estimators $\hat{\rho}_n^{(1)}(k)$ and $\hat{\rho}_n^{(2)}(k)$, for which a distributional representation of the type of the one in (10) holds, with constants (σ_1, b_1) and (σ_2, b_2) , $b_1, b_2 \neq 0$, respectively, both computed at their optimal levels, the AREFF of $\hat{\rho}_{n0}^{(1)}$ relatively to $\hat{\rho}_{n0}^{(2)}$ is*

$$\text{AREFF}_{1|2} \equiv \text{AREFF}_{\hat{\rho}_{n0}^{(1)}|\hat{\rho}_{n0}^{(2)}} := \sqrt{\frac{\text{LMSE}(\hat{\rho}_{n0}^{(2)})}{\text{LMSE}(\hat{\rho}_{n0}^{(1)})}} = \left(\left(\frac{\sigma_2}{\sigma_1} \right)^{-4\rho} \left| \frac{b_2}{b_1} \right|^{2(1-2\rho)} \right)^{\frac{1}{1-4\rho}}.$$

Let us think on $\text{AREFF}_{\text{FAGH}(\tau)|\text{FH}}$. Using the value τ_0^{FAGH} , in (18), we have $\text{AREFF}_{\text{FAGH}(\tau)|\text{FH}} > 1$ in a region $\tau \in \tau_0^{\text{FAGH}} \pm a_0^{\text{FAGH}}(\xi, \rho)$, with

$$a_0^{\text{FAGH}}(\xi, \rho) = \frac{6|2\xi - 1|}{|\rho|(1-3\rho)^2(3-\rho)(3-2\rho)} \left(\frac{(1-\rho)^{4(2-3\rho)}(1-2\rho)^{-2(1+\rho)}}{(2\rho^2 - 2\rho + 1)^{-2\rho}} \right)^{1/(2(1-2\rho))}. \quad (21)$$

The values $a_0^{\text{FAGH}}(\xi, \rho)$ are next presented at Figure 6, in the (ξ, ρ) -plane. See also Figure 7, where we present the regions of the (ξ, ρ) -plane where $\tau_0^{\text{FAGH}} - a_0^{\text{FAGH}}$ and $\tau_0^{\text{FAGH}} + a_0^{\text{FAGH}}$ are both non-negative, non-positive or with different signs. Note that for values of ρ close to zero and $\xi \neq 1/2$, the interval $(\tau_0^{\text{FAGH}} - a_0^{\text{FAGH}}, \tau_0^{\text{FAGH}} + a_0^{\text{FAGH}})$ covers the value $\tau = 0$, a value commonly used in practice for the region $|\rho| \leq 1$.

Similarly, let us now think on $\text{AREFF}_{\text{CG}(\tau)|\text{FH}}$. Using the value τ_0^{CG} , in (19), we have $\text{AREFF}_{\text{CG}(\tau)|\text{FH}} > 1$ in a region $\tau \in \tau_0^{\text{CG}} \pm a_0^{\text{CG}}(\xi, \rho)$, with

$$a_0^{\text{CG}}(\xi, \rho) = \frac{|2\xi - 1|(1-\rho)^2}{|\rho|(1-3\rho)^2} \left(\frac{(1-2\rho)^{-2(1+\rho)} 120^{-2\rho}}{(2-\rho)^{-4\rho}(3-2\rho)^{-4\rho}(4\rho^2 - 4\rho + 7)^{-2\rho}} \right)^{1/(2(1-2\rho))}. \quad (22)$$

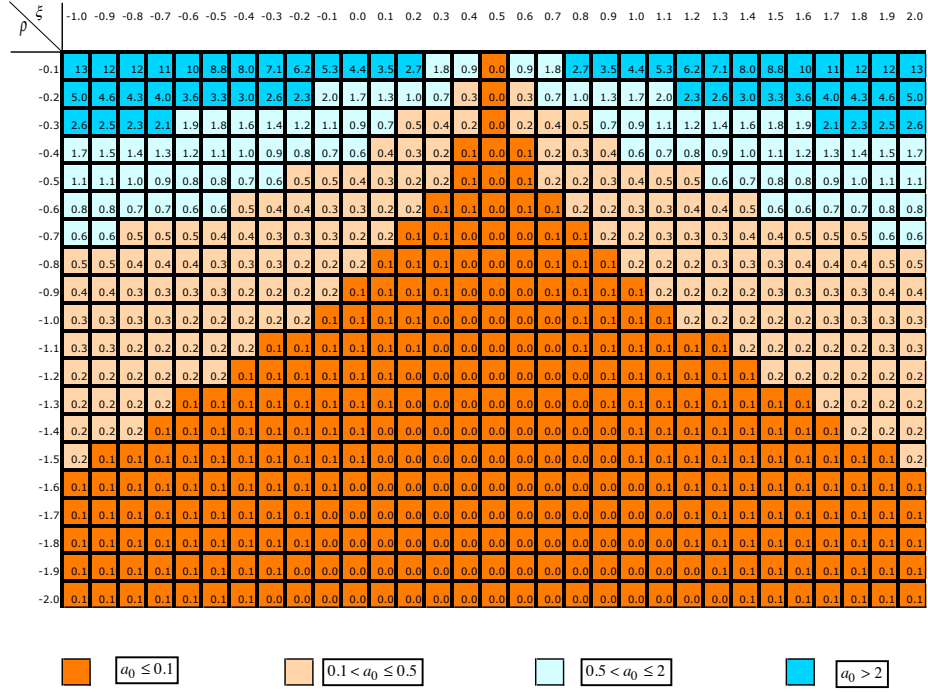


Figure 6: Values a_0 such that $\text{AREFF}_{\text{FAGH}(\tau)|_{\text{FH}}}^{\text{FAGH}} > 1$ for $\tau \in (\tau_0^{\text{FAGH}} - a_0, \tau_0^{\text{FAGH}} + a_0)$

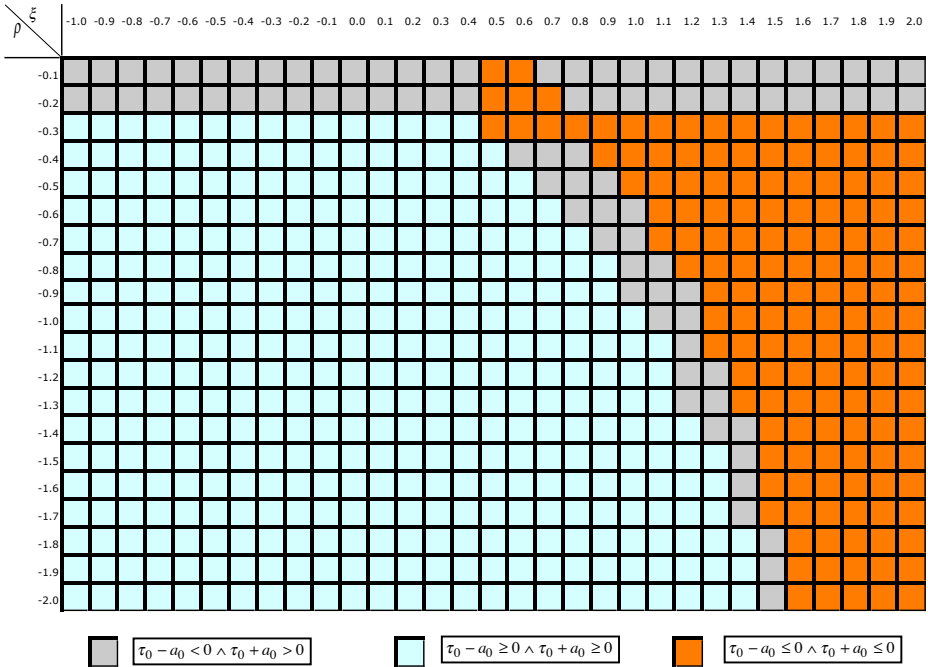


Figure 7: Signs of $\tau_0^{\text{FAGH}} - a_0$ and $\tau_0^{\text{FAGH}} + a_0$

The values $a_0^{\text{CG}}(\xi, \rho)$ are next presented at Figure 8, in the (ξ, ρ) -plane. See also Figure 9, where we present the signs of $\tau_0^{\text{CG}} - a_0^{\text{CG}}$ and $\tau_0^{\text{CG}} + a_0^{\text{CG}}$.

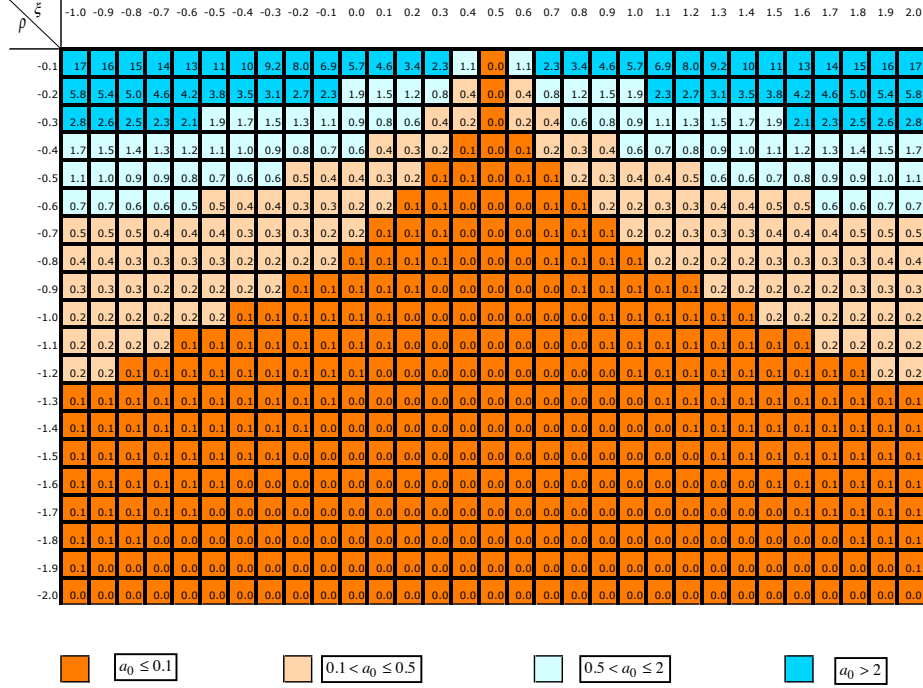


Figure 8: Values a_0 such that $\text{AREFF}_{\text{CG}(\tau)|\text{FH}} > 1$ for $\tau \in (\tau_0^{\text{CG}} - a_0, \tau_0^{\text{CG}} + a_0)$

For models with $\xi = 1/2$, $\hat{\rho}_{n0}^{\text{FH}}$ beats $\hat{\rho}_{n0}^{\text{FAGH}(\tau)}$ ($\hat{\rho}_{n0}^{\text{CG}(\tau)}$) unless we consider the tuning parameter τ_0^{FAGH} (τ_0^{CG}) in (18) ((19)), i.e. $\hat{\rho}_{n0}^{\text{FAGH}(\tau_0^{\text{FAGH}})}$ ($\hat{\rho}_{n0}^{\text{CG}(\tau_0^{\text{CG}})}$). For the asymptotic comparison between $\hat{\rho}_{n0}^{\text{FH}}$, $\hat{\rho}_{n0}^{\text{FAGH}(\tau_0^{\text{FAGH}})}$ and $\hat{\rho}_{n0}^{\text{CG}(\tau_0^{\text{CG}})}$ in the region $\xi = 1/2$ we should consider a fourth-order framework, a topic out of the scope of this article. The same comment applies to the asymptotic comparison at optimal levels, in the whole (ξ, ρ) -plane of the estimators $\hat{\rho}_{n0}^{\text{FAGH}(\tau_0^{\text{FAGH}})}$ and $\hat{\rho}_{n0}^{\text{CG}(\tau_0^{\text{CG}})}$.

As mentined before, $\sigma_{\text{FAGH}} > \sigma_{\text{FH}}$, for all $\rho < 0$. For any fixed $\tau \notin (\tau_0^{\text{FAGH}} - a_0^{\text{FAGH}}, \tau_0^{\text{FAGH}} + a_0^{\text{FAGH}})$, with τ_0^{FAGH} and a_0^{FAGH} given in (18) and (21), respectively, $\hat{\rho}_{n0}^{\text{FH}}$ beats $\hat{\rho}_{n0}^{\text{FAGH}(\tau)}$ in a wide region of the (ξ, ρ) -plane, only with the exclusion of values of ρ close to zero and a region around the line $b_{\text{FAGH}} = 0$. See the region in Figure 10 associated with $\tau = 0$, a value commonly used in practice for models with $|\rho| \leq 1$. A similar comment applies to the comparative behaviour of $\hat{\rho}_{n0}^{\text{FH}}$ and $\hat{\rho}_{n0}^{\text{CG}(\tau)}$ for any fixed $\tau \notin (\tau_0^{\text{CG}} - a_0^{\text{FAGH}}, \tau_0^{\text{CG}} + a_0^{\text{CG}})$, with τ_0^{CG} and a_0^{CG} given in (19) and (22),

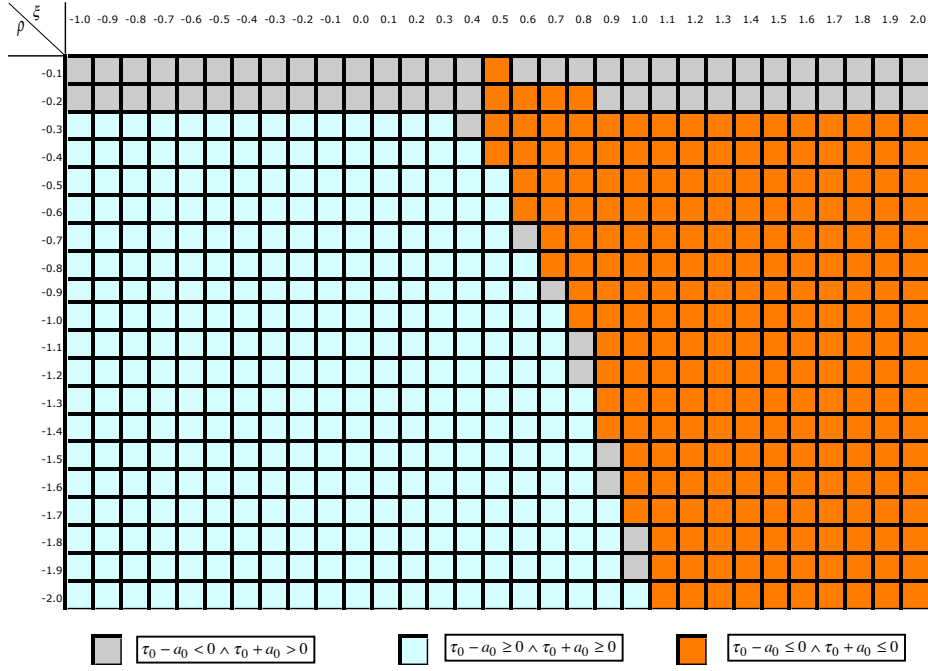


Figure 9: Values the signs of $\tau_0^{\text{CG}} - a_0$ and $\tau_0^{\text{CG}} + a_0$

respectively

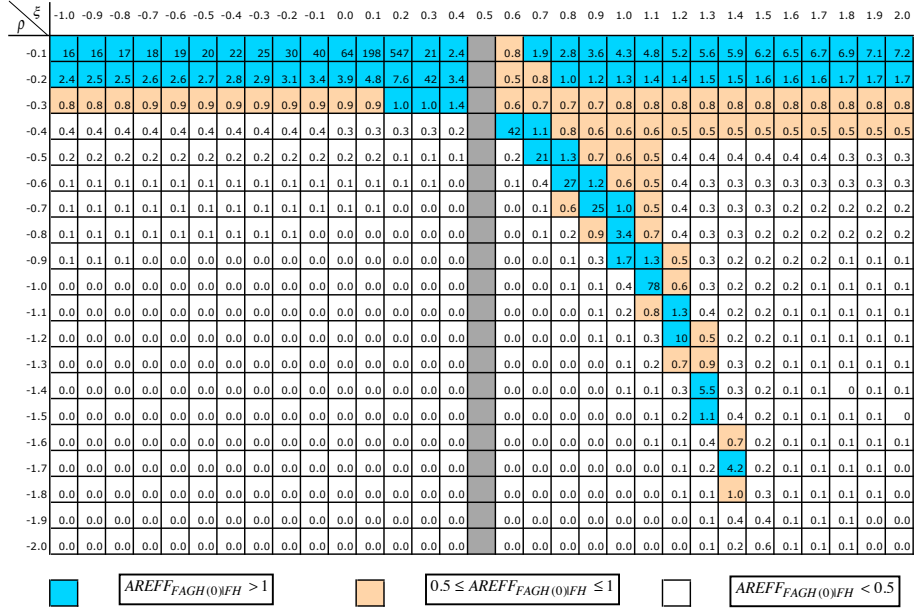


Figure 10: Values of $\text{AREFF}_{\text{FAGH}(\tau)|\text{FH}}$ for $\tau = 0$

For models with $\xi \neq 1/2$, $\hat{\rho}_{n0}^{\text{FAGH}(\tau_0^{\text{FAGH}})}$, as well as $\hat{\rho}_{n0}^{\text{CG}(\tau_0^{\text{CG}})}$, beat $\hat{\rho}_{n0}^{\text{FH}}$ for all (ξ, ρ) . Indeed, the same happens with $\hat{\rho}_{n0}^{\text{FAGH}(\tau)}$, $\tau \in (\tau_0^{\text{FAGH}} - a_0^{\text{FAGH}}, \tau_0^{\text{FAGH}} + a_0^{\text{FAGH}})$, with τ_0^{FAGH} and a_0^{FAGH} given in (18) and (21), respectively, and with $\hat{\rho}_{n0}^{\text{CG}(\tau)}$, $\tau \in (\tau_0^{\text{CG}} - a_0^{\text{CG}}, \tau_0^{\text{CG}} + a_0^{\text{CG}})$, with τ_0^{CG} and a_0^{CG} given in (19) and (22), respectively.

Regarding the asymptotic comparison at optimal levels of $\hat{\rho}_{n0}^{\text{FAGH}(\tau)}$ and $\hat{\rho}_{n0}^{\text{CG}(\tau)}$ for a fixed τ , we next present in Figures 11, 12, 13, 14 and 15, the values of $\text{AREFF}_{\text{FAGH}(\tau)|\text{G}(\tau)}$ for $\tau = -1, -0.5, 0, 0.5$ and 1 , respectively. Due to the fact that for the most common models in practice $\xi > 0.5$, we can say that for a fixed τ , $\hat{\rho}_{n0}^{\text{FAGH}(\tau)}$ beats $\hat{\rho}_{n0}^{\text{CG}(\tau)}$, at optimal levels, in a wide region of the (ξ, ρ) -plane.

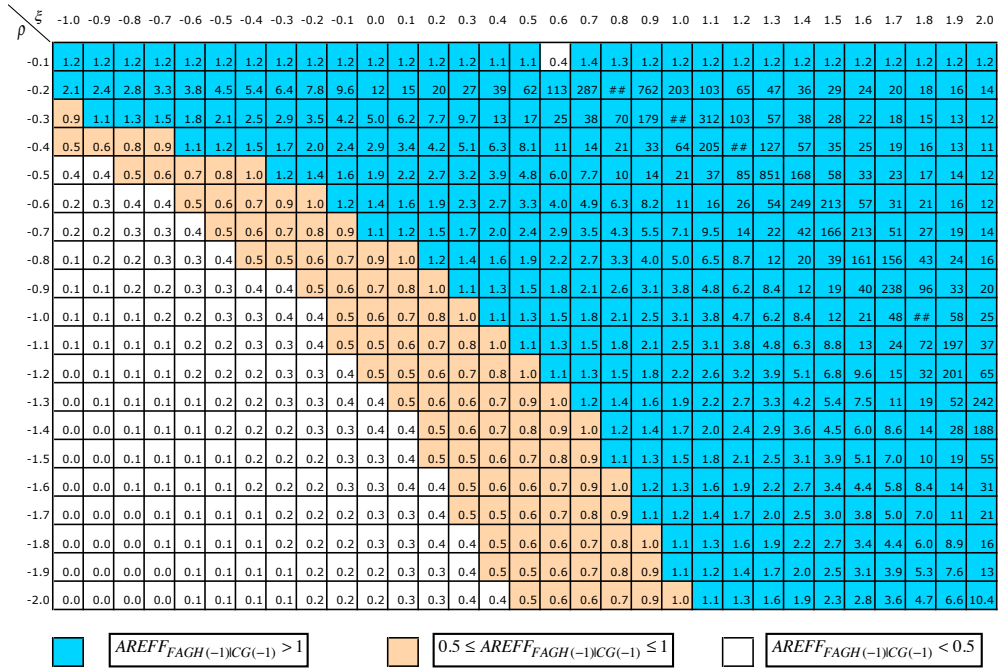


Figure 11: Values of $\text{AREFF}_{\text{FAGH}(\tau)|\text{CG}}$ for $\tau = -1$

In practice, it is however sensible the adequate choice of a value of τ close to either τ_0^{FAGH} or τ_0^{CG} . To make such a choice we can use any heuristic algorithm based on sample stability, like the one used in Caeiro and Gomes (2012b), similar in spirit to the ones in Figueiredo *et al.* (2012) for value-at-risk-estimation and in Gomes *et al.* (2011a,b) for estimation of the extreme value index.

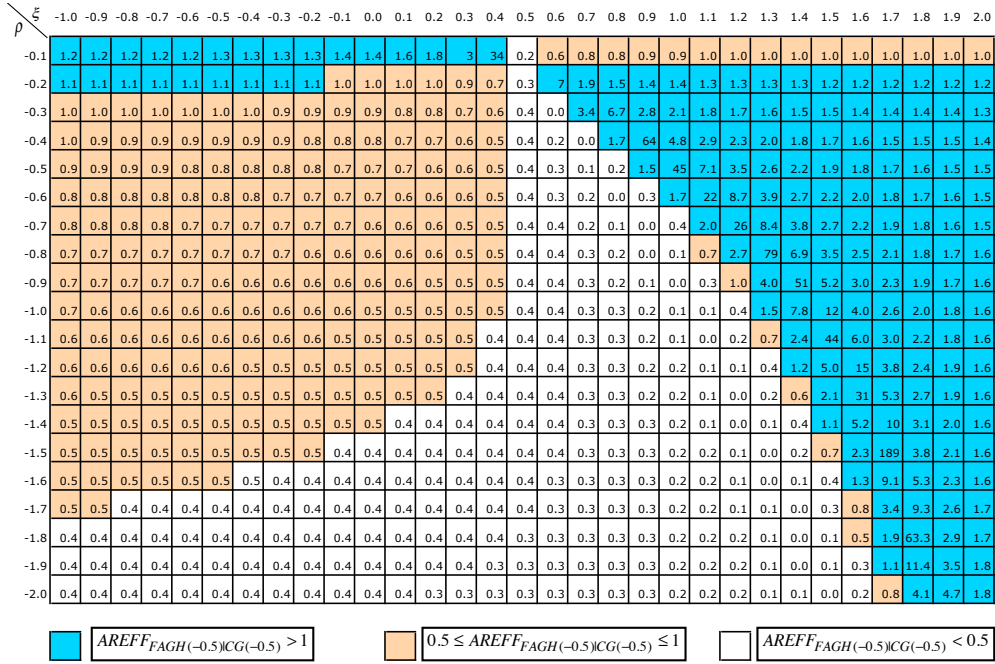


Figure 12: Values of $AREFF_{FAGH(\tau)|CG}$ for $\tau = -0.5$

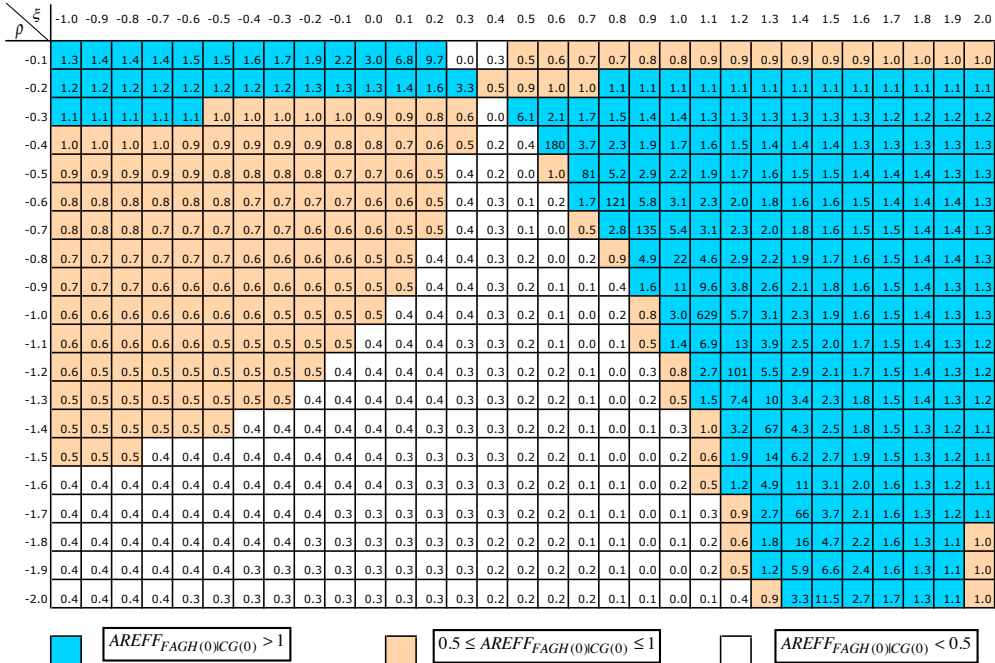


Figure 13: Values of $AREFF_{FAGH(\tau)|CG}$ for $\tau = 0$

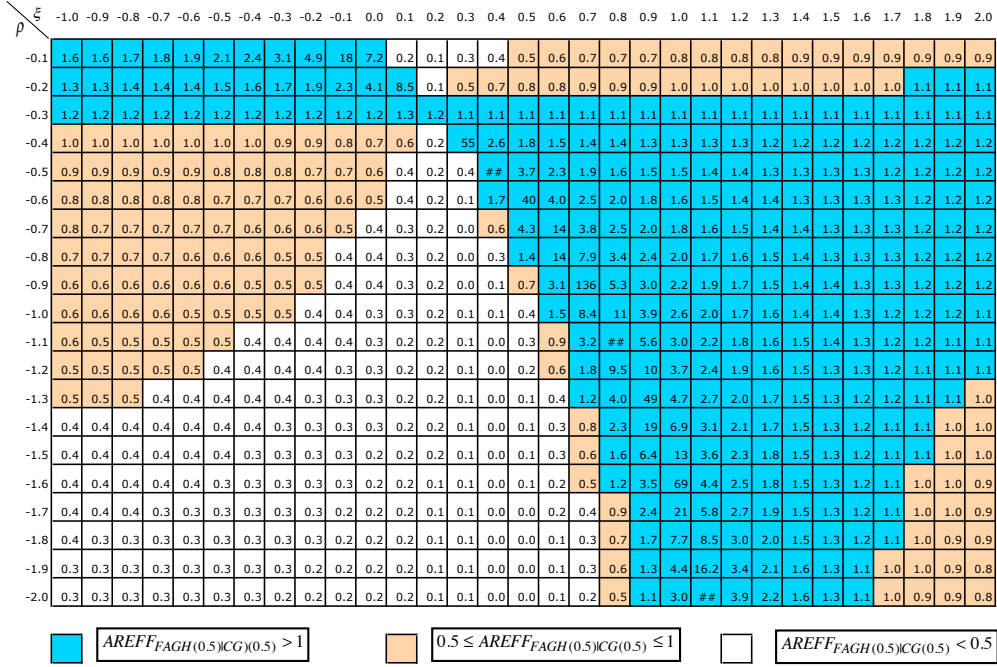


Figure 14: Values of $AREFF_{FAGH(\tau)|CG}$ for $\tau = 0.5$

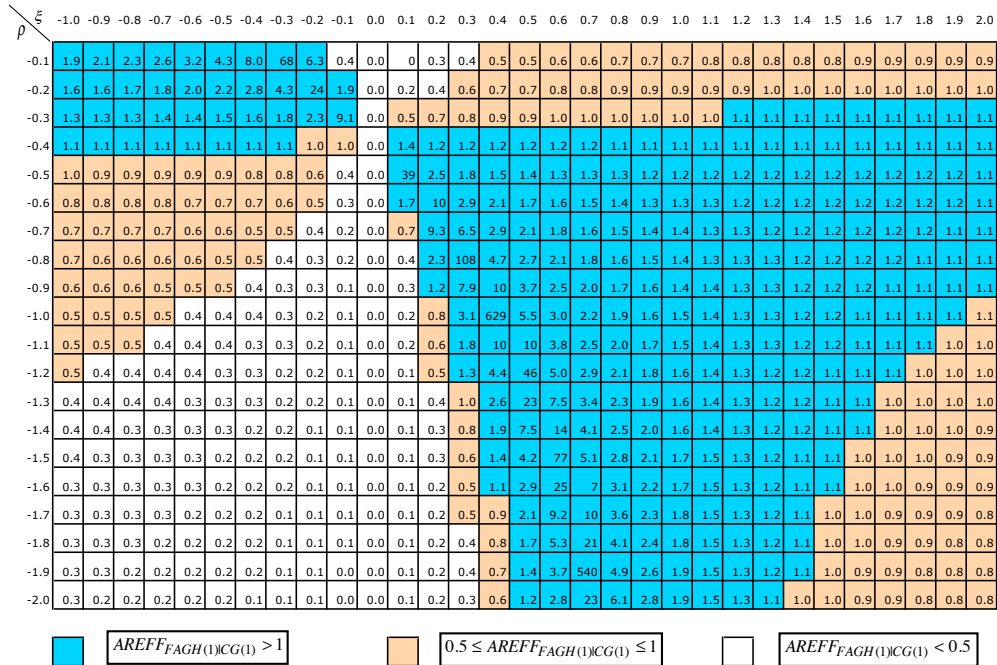


Figure 15: Values of $AREFF_{FAGH(\tau)|CG}$ for $\tau = 1$

References

- [1] Bingham, N., Goldie, C.M., and Teugels, J.L. (1987). *Regular Variation*. Cambridge Univ. Press, Cambridge.
- [2] Caeiro, F., and Gomes, M.I. (2006). A new class of estimators of a “scale” second order parameter. *Extremes* **9**, 193–211.
- [3] Caeiro, F., and Gomes, M.I. (2008). Minimum-variance reduced-bias tail index and high quantile estimation. *Revstat* **6**:1, 1–20.
- [4] Caeiro, F. and Gomes, M.I. (2011). Asymptotic comparison at optimal levels of reduced-bias extreme value index estimators. *Statistica Neerlandica* **65**:4, 462–488.
- [5] Caeiro, F. and Gomes, M.I. (2012a). *A Semi-Parametric Estimator of a Shape Second Order Parameter*, Notas e Comunicações CEAUL 07/2012, submitted.
- [6] Caeiro, F. and Gomes, M.I. (2012b). *Bias Reduction in the Estimation of a Shape Second-Order Parameter of a Heavy Right Tail Model*, Preprint, CMA 22-2012, submitted.
- [7] Caeiro, F., Gomes, M.I. and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework. *Commun. Statist.—Theory and Methods* **38**:7, 1019–1040.
- [8] Ciuperca, G. and Mercadier, C. (2010). Semi-parametric estimation for heavy tailed distributions. *Extremes* **13**:1, 55–87.
- [9] Feuerverger, A. and Hall, P. (1999) Estimating a tail exponent by modelling departure from a Pareto distribution. *Ann. Statist.* **27**, 760–781.
- [10] Figueiredo, F., Gomes, M.I., Henriques-Rodrigues, L. and Miranda, C. (2012). A computational study of a quasi-PORT methodology for VaR based on second-order reduced-bias estimation. *J. Statist. Comput. and Simul.* **82**:4, 587–602.

- [11] 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.
- [12] 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.
- [13] 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.
- [14] Goegebeur, Y., Beirlant, J. and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *J. Statist. Planning and Inference* **140**:9, 2632–2652.
- [15] Gomes, M.I. and Henriques-Rodrigues, L. (2010). Comparison at optimal levels of classical tail index estimators: a challenge for reduced-bias estimation? *Discussiones Mathematica: Probability and Statistics* **30**:1, 35–51.
- [16] Gomes, M.I. and Martins, M.J. (2001). Generalizations of the Hill estimator — asymptotic versus finite sample behaviour. *J. Statistical Planning and Inference* **93**, 161–180.
- [17] Gomes, M.I. and Neves, C. (2008). Asymptotic comparison of the mixed moment and classical extreme value index estimators. *Statistics and Probability Letters* **78**:6, 643–653.
- [18] Gomes, M.I., Miranda, C. and Pereira, H. (2005). Revisiting the role of the Jackknife methodology in the estimation of a positive extreme value index. *Communications in Statistics—Theory and Methods* **34**, 1–20.
- [19] Gomes, M.I., Martins, M.J. and Neves, M.M. (2007a). Improving second order reduced-bias extreme value index estimation. *Revstat* **5**:2, 177–207.
- [20] Gomes, M.I., Miranda, C. and Viseu, C. (2007b). Reduced-bias tail index estimation and the Jackknife methodology. *Statistica Neerlandica* **61**:2, 243–270.

- [21] Gomes, M.I., Henriques-Rodrigues, L. and Miranda, C. (2011a). Reduced-bias location-invariant extreme value index estimation: a simulation study. *Comm.Statist.–Simul. and Comput.* **40**:3, 424–447.
- [22] 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. and Simul.*, in press. DOI:10.1080/00949655.2011.652113
- [23] Haan, L. de and Peng, L. (1998). Comparison of extreme value index estimators. *Statistica Neerlandica* **52**, 60–70.
- [24] Hall, P. (1982). On estimating the endpoint of a distribution. *Ann. Statist.* **10**, 556-568.
- [25] Hall, P., and Welsh, A.W. (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.* **13**, 331–341.