



Weighted Moment Estimators for the Second Order Scale Parameter

Tertius de Wet, Yuri Goegebeur, Armelle Guillou

► To cite this version:

Tertius de Wet, Yuri Goegebeur, Armelle Guillou. Weighted Moment Estimators for the Second Order Scale Parameter. Methodology and Computing in Applied Probability, 2012, 14 (3), 10.1007/s11009-011-9263-6 . hal-01311745

HAL Id: hal-01311745

<https://hal.science/hal-01311745>

Submitted on 11 May 2016

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

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

Weighted moment estimators for the second order scale parameter

Tertius de Wet ^{*}
Yuri Goegebeur [†]
Armelle Guillou [‡]

September 14, 2011

Abstract

We consider the estimation of the scale parameter appearing in the second order condition when the distribution underlying the data is of Pareto-type. Inspired by the work of Goegebeur, Beirlant and de Wet (2010) on the estimation of the second order rate parameter, we introduce a flexible class of estimators for the second order scale parameter, which has weighted sums of scaled log spacings of successive order statistics as basic building blocks. Under the second order condition, some conditions on the weight functions, and for appropriately chosen sequences of intermediate order statistics, we establish the consistency of our class of estimators. Asymptotic normality is achieved under a further condition on the tail function $1 - F$, the so-called third order condition. As the proposed estimator depends on the second order rate parameter, we also examine the effect of replacing the latter by a consistent estimator. The asymptotic performance of some specific examples of our proposed class of estimators is illustrated numerically, and their finite sample behavior is examined by a small simulation experiment.

Keywords: Extreme value statistics, Pareto-type model, second order scale parameter, weighted estimator.

1 Introduction

When interest is in estimating parameters related to the far tail of a distribution function, such as extreme quantiles or small exceedance probabilities, one has to extend the empirical distribution function beyond the available data. Extreme value theory studies the behavior of the largest observations in a sample and provides laws governing these values, and as such forms the natural framework for carrying out this extrapolation step. We refer to the seminal works by Fréchet (1927) and Fisher and Tippett (1928). In this paper we deal with an estimation problem

^{*}Department of Statistics and Actuarial Science, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa (email: tdewet@sun.ac.za). This author's research was partially supported by grants from the National Research Foundation and ESKOM.

[†]Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark (email: yuri.goegebeur@stat.sdu.dk). This author's research was supported by a grant from the Danish Natural Science Research Council.

[‡]Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France (email: armelle.guillou@math.unistra.fr).

within the class of heavy-tailed or Pareto-type distributions, which constitute the max-domain of attraction of the Fréchet distribution. Such models find important practical applications in areas such as non-life insurance, finance, telecommunications and geology, to name but a few.

A distribution is said to be of Pareto-type if for some $\gamma > 0$ its survival function is of the form:

$$1 - F(x) = x^{-1/\gamma} \ell_F(x), \quad x > 0, \quad (1)$$

where ℓ_F denotes a slowly varying function at infinity, i.e.

$$\frac{\ell_F(\lambda x)}{\ell_F(x)} \rightarrow 1 \text{ as } x \rightarrow \infty \text{ for all } \lambda > 0. \quad (2)$$

The parameter γ is called the extreme value index, and clearly governs the tail behavior, with larger values indicating heavier tails. The Pareto-type model can also be stated in an equivalent way in terms of the tail quantile function U , where $U(x) := \inf\{y : F(y) \geq 1 - 1/x\}$, $x > 1$, as follows

$$U(x) = x^\gamma \ell_U(x), \quad (3)$$

with ℓ_U again a slowly varying function at infinity (Gnedenko, 1943).

A central topic in the analysis of Pareto-type distributions is the estimation of the extreme value index γ . Several estimators for this parameter have been proposed in the statistical literature, and their limiting distributions established, usually under a second order condition on the tail behavior.

Second order condition ($\mathcal{R}$) *There exists a positive real parameter γ , a negative real parameter ρ and a function b with $b(t) \rightarrow 0$ for $t \rightarrow \infty$, of constant sign for large values of t , such that*

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

We refer to Bingham *et al.* (1987) and de Haan and Ferreira (2006) for further details. As in Caeiro and Gomes (2006) and Caeiro *et al.* (2009), we assume that $(\mathcal{R})$ holds with $b(t) = \gamma D t^\rho$, $D \neq 0$. Note that as such the parameter ρ controls the rate of convergence in the first order framework given in (3). When $|\rho|$ is small this convergence is slow and the estimation of tail parameters is typically difficult in practice. Condition $(\mathcal{R})$ is not too restrictive; for instance the important Hall-Welsh class of Pareto-type models (Hall and Welsh, 1985) for which the tail quantile function is of the form

$$U(x) = C x^\gamma \left(1 + \frac{\gamma D}{\rho} x^\rho + o(x^\rho) \right),$$

satisfies $(\mathcal{R})$. The estimation of the second order parameters ρ and D are challenging estimation problems. The estimation of the second order rate parameter ρ has received considerable

attention in the recent extreme value literature; see Gomes *et al.* (2002), Fraga Alves *et al.* (2003), Ciuperca and Mercadier (2010) and Goegebeur *et al.* (2010) for examples of estimators that show good practical performance. The estimation of the second order scale parameter D is less well studied. Gomes and Martins (2002) introduced a maximum likelihood estimator for D derived from the representation of Beirlant *et al.* (1999) and Feuerverger and Hall (1999) for scaled log-spacings of successive order statistics, whereas Gomes *et al.* (2010) applied a quasi likelihood estimation method to weighted log-excesses over a high threshold. Caeiro and Gomes (2006) studied an estimator based on generalized Hill statistics, the type of statistics used in e.g. Fraga Alves *et al.* (2003) for the construction of estimators for ρ . In this paper we will elaborate on the work Goegebeur *et al.* (2010) performed on the estimation of the second order rate parameter ρ . In particular, we will consider weighted sums of scaled log spacings of successive order statistics, and based on their asymptotic distributional representations introduce a flexible class of estimators for D . Adequate estimation of the second order scale parameter D has practical relevance for bias-corrected estimation of the extreme value index γ . Indeed, recent research on tail index estimation has focused on the development of improved estimators for γ by explicitly estimating the dominant term of the asymptotic bias. This bias depends on D and ρ , so an a priori estimation of these parameters is necessary to obtain the bias reduction. We refer to Beirlant *et al.* (1999), Feuerverger and Hall (1999), Gomes and Martins (2002), Caeiro *et al.* (2005), and more recently Gomes *et al.* (2008) for examples of such bias-corrected estimators.

The remainder of this paper is organized as follows. In the next section we introduce the weighted moment estimator for the second order scale parameter and establish its consistency under the second order condition $(\mathcal{R})$, and some conditions on the weight function and on the number of upper order statistics used in the estimation. We also show that our estimator, when appropriately normalized, converges in distribution to a normal random variable, if we invoke a third order condition on the tail behavior, some further conditions on the weight function and on the number of upper order statistics the estimator is based upon. The asymptotic and finite sample behavior of estimators obtained for some specific examples of weight functions are illustrated numerically in Section 3. The proofs of the main results are deferred to the appendix.

2 Asymptotic properties

2.1 Consistency

In this section we introduce our estimator for D and establish its consistency. Let $X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) random variables from a distribution function satisfying (1) and denote by $X_{1,n} \leq \dots \leq X_{n,n}$ the associated order statistics. The basic building blocks for the estimator are the kernel statistics considered in Goegebeur *et al.* (2008, 2010):

$$T_{n,k}(K) := \frac{1}{k} \sum_{j=1}^k K\left(\frac{j}{k+1}\right) Z_j,$$

where K is a non-negative weight function and $Z_j := j(\ln X_{n-j+1,n} - \ln X_{n-j,n})$, $j = 1, \dots, k$.

In order to obtain the asymptotic distributional expansion of $T_{n,k}(K)$ we have to introduce some conditions on the weight function K . Define $\ln^+(1/u) := \max\{\ln(1/u), 1\}$ with $u \in (0, 1)$, $\mu(K) := \int_0^1 K(u)du$ and $\sigma^2(K) := \int_0^1 K^2(u)du$.

Condition (K) Let K be a function defined on $(0, 1)$ such that

(i) $K(t) = \frac{1}{t} \int_0^t u(v)dv$ for some function u satisfying $\left| (k+1) \int_{(j-1)/(k+1)}^{j/(k+1)} u(t)dt \right| \leq f(\frac{j}{k+1})$ for some positive continuous function f defined on $(0, 1)$ such that $\int_0^1 \ln^+(1/w)f(w)dw < \infty$,

(ii) $\sigma^2(K) < \infty$,

(iii) $\frac{1}{k} \sum_{j=1}^k K\left(\frac{j}{k+1}\right) = \mu(K) + o(1/\sqrt{k})$ for $k \rightarrow \infty$,

(iv) $\max_{i \in \{1, \dots, k\}} K\left(\frac{i}{k+1}\right) = o(\sqrt{k})$ for $k \rightarrow \infty$.

Let $K_1, \dots, K_8$ be kernel functions, $\mathbb{K}^{(1)} := (K_1, \dots, K_4)$, $\mathbb{K}^{(2)} := (K_5, \dots, K_8)$, $\mathbb{K}^{(1,2)} := (\mathbb{K}^{(1)}, \mathbb{K}^{(2)})$,

$$I(K_i, \rho) := \int_0^1 K_i(u)u^{-\rho}du, \quad \bar{I}(K_i, \rho) := \frac{I(K_i, \rho)}{\mu(K_i)}, \quad \bar{I}^{(a)}(K_i, K_j, \rho) := \bar{I}^a(K_i, \rho) - \bar{I}^a(K_j, \rho),$$

where $a = 1, 2$ and $i, j \in \{1, 2, \dots, 8\}$. We assume that $\bar{I}^{(1)}(K_1, K_2, \rho)$, $\bar{I}^{(1)}(K_3, K_4, \rho)$, $\bar{I}^{(1)}(K_5, K_6, \rho)$ and $\bar{I}^{(1)}(K_7, K_8, \rho)$ are well defined and non-zero.

Consider the statistic

$$\Psi_{n,k}(\mathbb{K}^{(1)}, \alpha_1, \theta) := \frac{\left[\left(\frac{T_{n,k}(K_1)}{\mu(K_1)} \right)^{\alpha_1} - \left(\frac{T_{n,k}(K_2)}{\mu(K_2)} \right)^{\alpha_1} \right]^\theta}{\left(\frac{T_{n,k}(K_3)}{\mu(K_3)} \right)^{\alpha_1 \theta} - \left(\frac{T_{n,k}(K_4)}{\mu(K_4)} \right)^{\alpha_1 \theta}},$$

where $\alpha_1 \in \mathbb{R}$ and $\theta \in \mathbb{Z} \setminus \{0, 1\}$. Under conditions $(\mathcal{R})$ and $(\mathcal{K})$ we have, for $k, n \rightarrow \infty$ such that $k/n \rightarrow 0$ and $\sqrt{k}b(n/k) \rightarrow \infty$, the following convergence

$$\left(\frac{k}{n} \right)^{\rho(\theta-1)} \Psi_{n,k}(\mathbb{K}^{(1)}, \alpha_1, \theta) \xrightarrow{\mathbb{P}} \frac{(\alpha_1 D)^{\theta-1} [\bar{I}^{(1)}(K_1, K_2, \rho)]^\theta}{\theta \bar{I}^{(1)}(K_3, K_4, \rho)}. \quad (4)$$

Similarly, under the same conditions

$$\left(\frac{k}{n} \right)^{\rho(\theta-2)} \Psi_{n,k}(\mathbb{K}^{(2)}, \alpha_2, \theta-1) \xrightarrow{\mathbb{P}} \frac{(\alpha_2 D)^{\theta-2} [\bar{I}^{(1)}(K_5, K_6, \rho)]^{\theta-1}}{(\theta-1) \bar{I}^{(1)}(K_7, K_8, \rho)}, \quad (5)$$

where $\alpha_2 \in \mathbb{R}$. Equations (4) and (5) motivate then the following estimator for D :

$$\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) := c(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) \left(\frac{k}{n} \right)^\rho \Lambda_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta),$$

where

$$c(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) := \frac{\theta \alpha_2^{\theta-2} \bar{I}^{(1)}(K_3, K_4, \rho) [\bar{I}^{(1)}(K_5, K_6, \rho)]^{\theta-1}}{(\theta-1) \alpha_1^{\theta-1} \bar{I}^{(1)}(K_7, K_8, \rho) [\bar{I}^{(1)}(K_1, K_2, \rho)]^\theta},$$

$$\Lambda_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta) := \frac{\Psi_{n,k}(\mathbb{K}^{(1)}, \alpha_1, \theta)}{\Psi_{n,k}(\mathbb{K}^{(2)}, \alpha_2, \theta-1)}.$$

Note that the proposed estimator for D depends on ρ , the rate parameter appearing in the second order condition, which is typically unknown. If this is the case then one replaces ρ by a consistent estimator $\hat{\rho}_k$.

Theorem 1 *Let $X_1, \dots, X_n$ be i.i.d. random variables according to a distribution satisfying $(\mathcal{R})$. Let $K_1, \dots, K_8$ satisfy condition $(\mathcal{K})$, and suppose $\bar{I}^{(1)}(K_1, K_2, \rho)$, $\bar{I}^{(1)}(K_3, K_4, \rho)$, $\bar{I}^{(1)}(K_5, K_6, \rho)$ and $\bar{I}^{(1)}(K_7, K_8, \rho)$ are well defined and nonzero. Then, if $k, n \rightarrow \infty$ such that $k/n \rightarrow 0$ and $\sqrt{kb}(n/k) \rightarrow \infty$ we have $\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) \xrightarrow{\mathbb{P}} D$. Further, under the same conditions, $\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \hat{\rho}_k) \xrightarrow{\mathbb{P}} D$ for any consistent estimator $\hat{\rho}_k$ for ρ that satisfies $(\hat{\rho}_k - \rho) \ln n/k = o_{\mathbb{P}}(1)$.*

Goegebeur *et al.* (2010) introduced a very flexible class of consistent estimators for the second order rate parameter ρ . Based on the asymptotic expansion of $T_{n,k}(K)$ they proposed the following ratio

$$\Delta_{n,k}(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l) := \frac{\tilde{\Delta}_{n,k}(\mathbb{K}^{(1)}, \omega_1, \omega_1 + l)}{\tilde{\Delta}_{n,k}(\mathbb{K}^{(2)}, \omega_2, \omega_2 + l)}$$

where ω_1, ω_2 and l are positive tuning parameters, and with

$$\tilde{\Delta}_{n,k}(\mathbb{K}^{(1)}, \omega_1, \omega_2) := \frac{\left(\frac{T_{n,k}(K_1)}{\mu(K_1)}\right)^{\omega_1} - \left(\frac{T_{n,k}(K_2)}{\mu(K_2)}\right)^{\omega_1}}{\left(\frac{T_{n,k}(K_3)}{\mu(K_3)}\right)^{\omega_2} - \left(\frac{T_{n,k}(K_4)}{\mu(K_4)}\right)^{\omega_2}}.$$

Under the conditions of Theorem 1 it was shown that $\Delta_{n,k}(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l) \xrightarrow{\mathbb{P}} \Delta(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \rho)$, a function depending on the unknown ρ only, which then leads to the estimator $\hat{\rho}_k(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l) := \Delta^{-1}(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \Delta_{n,k}(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l))$, provided the function $\rho \mapsto \Delta(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \rho)$ is bijective. If Δ^{-1} is continuous then the estimator $\hat{\rho}_k(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l)$ is consistent for ρ . Other consistent estimators for ρ can be found in Gomes *et al.* (2002), Fraga Alves *et al.* (2003), and Ciuperca and Mercadier (2010).

2.2 Asymptotic normality

In this section we will establish the asymptotic normality of our estimator for D , when appropriately normalized. Similar to the estimation of the second order rate parameter ρ , we need a third order condition on the tail behavior and some further conditions on the weight functions. We start by introducing the third order condition on the tail of F .

Third order condition ($\tilde{\mathcal{R}}$) *There exists a positive real parameter γ , negative real parameters ρ and β , functions b and $\tilde{b}$ with $b(t) \rightarrow 0$ and $\tilde{b}(t) \rightarrow 0$ for $t \rightarrow \infty$, both of constant sign for large values of t , such that*

$$\lim_{t \rightarrow \infty} \frac{\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{b(t)} - \frac{x^\rho - 1}{\rho}}{\tilde{b}(t)} = \frac{x^{\rho+\beta} - 1}{\rho + \beta}, \quad \forall x > 0.$$

We assume here that $\tilde{b}(t) = \xi t^\beta$, $\xi \neq 0$. The parameter β thus determines the rate of convergence in the second order framework given in $(\mathcal{R})$, and hence plays a role that is analogous to that of the second order rate parameter ρ . Recently, Goegebeur and de Wet (2011) used statistics of the type $T_{n,k}(K)$ to introduce a class of estimators for β and studied its asymptotic properties. The third order condition is not too restrictive and is satisfied for a large subclass of the Hall-Welsh class of Pareto-type models (Hall and Welsh, 1985). In Table 1 we provide some examples of Pareto-type models satisfying $(\tilde{\mathcal{R}})$, where we derive for each of them the parameter γ , and the functions $b(x)$ and $\tilde{b}(x)$.

Concerning the kernel function K we assume the following:

Condition ($\tilde{\mathcal{K}}$) *Assume $(\mathcal{K})$ with (i) replaced by*

(i') $K(t) = \frac{1}{t} \int_0^t u(v) dv$ for some function u satisfying $\left| (k+1) \int_{(j-1)/(k+1)}^{j/(k+1)} u(t) dt \right| \leq f(\frac{j}{k+1})$ for some positive continuous and integrable function f defined on $(0, 1)$,

and the following two extra conditions

(v) $\frac{1}{k} \sum_{j=1}^k K\left(\frac{j}{k+1}\right) \left(\frac{j}{k+1}\right)^{-\rho} = I_1(K, \rho) + o(1/\sqrt{k})$ for $k \rightarrow \infty$,

(vi) $\int_0^1 K(u) u^{|\rho|-1-\varepsilon} du < \infty$ for some $\varepsilon > 0$.

Conditions $(\mathcal{K})$ and $(\tilde{\mathcal{K}})$ are satisfied by the class of weight functions given by

$$K(u; \tau, \delta) := (-\ln u)^\tau u^\delta, \quad u \in (0, 1); \quad \tau, \delta \geq 0. \quad (6)$$

We refer to Lemma 3 for the details about the verification of $(\mathcal{K})$ – $(\tilde{\mathcal{K}})$. The class of weight functions (6) is very flexible and contains for instance the power weight function $\mathbb{P}_\delta := K(\cdot; 0, \delta)$ and the log weight function $\mathbb{L}_\tau := K(\cdot; \tau, 0)$, considered in Goegebeur *et al.* (2010) in the framework of the estimation of ρ .

As will be clear from Theorem 2 given below, if we estimate D and ρ by means of the same sequence for k , then the asymptotic behavior of $\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \hat{\rho}_k)$ is determined by that of $\hat{\rho}_k$. In the sequel we will concentrate on the estimators for ρ introduced in Goegebeur *et al.* (2010), though similar results can be obtained with the ρ estimators of Gomes *et al.* (2002), Fraga Alves *et al.* (2003) and Ciuperca and Mercadier (2010). Goegebeur *et al.* (2010) showed that if U satisfies $(\tilde{\mathcal{R}})$, the weight functions in $\mathbb{K}^{(1,2)}$ satisfy $(\tilde{\mathcal{K}})$, and Δ is bijective and differentiable

in ρ , then for $k, n \rightarrow \infty$ such that $k/n \rightarrow 0$, $\sqrt{kb}(n/k) \rightarrow \infty$, $\sqrt{kb}(n/k)\tilde{b}(n/k) \rightarrow \lambda_1$ and $\sqrt{kb^2}(n/k) \rightarrow \lambda_2$ one obtains the following convergence in distribution

$$\begin{aligned} & \sqrt{kb}(n/k) \left[\hat{\rho}_{n,k}(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l) - \rho \right] \\ & \xrightarrow{\mathcal{D}} N \left(\lambda_1 \beta \tilde{c}_1(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho, \beta) + \lambda_2 \tilde{c}_2(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho), \tilde{v}^2(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho) \right). \end{aligned}$$

We refer to Goegebeur *et al.* (2010) for details about $\tilde{c}_1(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho, \beta)$, $\tilde{c}_2(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho)$ and $\tilde{v}^2(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho)$.

Let

$$\begin{aligned} \bar{\sigma}^2(K_i) &:= \frac{\sigma^2(K_i)}{\mu^2(K_i)}, \\ N_k(K_i, K_j) &:= \bar{\sigma}(K_i)N_k(K_i) - \bar{\sigma}(K_j)N_k(K_j), \\ N_k(\mathbb{K}^{(1,2)}, \theta, \rho) &:= \gamma D \left[\frac{\theta N_k(K_1, K_2)}{\bar{I}^{(1)}(K_1, K_2, \rho)} - \frac{N_k(K_3, K_4)}{\bar{I}^{(1)}(K_3, K_4, \rho)} - \frac{(\theta - 1)N_k(K_5, K_6)}{\bar{I}^{(1)}(K_5, K_6, \rho)} + \frac{N_k(K_7, K_8)}{\bar{I}^{(1)}(K_7, K_8, \rho)} \right], \\ u_D(\mathbb{K}^{(1,2)}, \theta, \rho, \beta) &:= D \left[\frac{\theta \bar{I}^{(1)}(K_1, K_2, \rho + \beta)}{\bar{I}^{(1)}(K_1, K_2, \rho)} - \frac{\bar{I}^{(1)}(K_3, K_4, \rho + \beta)}{\bar{I}^{(1)}(K_3, K_4, \rho)} \right. \\ &\quad \left. - \frac{(\theta - 1)\bar{I}^{(1)}(K_5, K_6, \rho + \beta)}{\bar{I}^{(1)}(K_5, K_6, \rho)} + \frac{\bar{I}^{(1)}(K_7, K_8, \rho + \beta)}{\bar{I}^{(1)}(K_7, K_8, \rho)} \right], \\ v_D(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) &:= \frac{D}{2\gamma} \left[\frac{\theta(\alpha_1 - 1)\bar{I}^{(2)}(K_1, K_2, \rho)}{\bar{I}^{(1)}(K_1, K_2, \rho)} - \frac{(\alpha_1\theta - 1)\bar{I}^{(2)}(K_3, K_4, \rho)}{\bar{I}^{(1)}(K_3, K_4, \rho)} \right. \\ &\quad \left. - \frac{(\theta - 1)(\alpha_2 - 1)\bar{I}^{(2)}(K_5, K_6, \rho)}{\bar{I}^{(1)}(K_5, K_6, \rho)} + \frac{(\alpha_2(\theta - 1) - 1)\bar{I}^{(2)}(K_7, K_8, \rho)}{\bar{I}^{(1)}(K_7, K_8, \rho)} \right], \\ w_D^2(\mathbb{K}^{(1,2)}, \theta, \rho) &:= \text{Avar}(N_k(\mathbb{K}^{(1,2)}, \theta, \rho)). \end{aligned}$$

Theorem 2 Let $X_1, \dots, X_n$ be i.i.d. random variables according to a distribution satisfying $(\tilde{\mathcal{R}})$. If the kernel functions $K_1, \dots, K_8$ satisfy assumption $(\tilde{\mathcal{K}})$ and are such that $\bar{I}_1^{(1)}(K_1, K_2, \rho)$, $\bar{I}_1^{(1)}(K_3, K_4, \rho)$, $\bar{I}_1^{(1)}(K_5, K_6, \rho)$, $\bar{I}_1^{(1)}(K_7, K_8, \rho)$ are well defined and nonzero, then for $k, n \rightarrow \infty$ such that $k/n \rightarrow 0$, $\sqrt{kb}(n/k) \rightarrow \infty$, $\sqrt{kb}(n/k)\tilde{b}(n/k) \rightarrow \lambda_1$ and $\sqrt{kb^2}(n/k) \rightarrow \lambda_2$ we have

$$\begin{aligned} & \sqrt{kb}(n/k) \left[\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) - D \right] \\ & \xrightarrow{\mathcal{D}} N \left(\lambda_1 u_D(\mathbb{K}^{(1,2)}, \theta, \rho, \beta) + \lambda_2 v_D(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho), w_D^2(\mathbb{K}^{(1,2)}, \theta, \rho) \right). \end{aligned} \quad (7)$$

The same result continues to hold when ρ in $\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho)$ is replaced by an external estimator $\hat{\rho}_{\check{k}}$ which is such that $\hat{\rho}_{\check{k}} - \rho = O_{\mathbb{P}}(1/\sqrt{\check{k}b}(n/\check{k}))$, when $\sqrt{\check{k}b}(n/\check{k}) \rightarrow \infty$, provided

$$\frac{\sqrt{\check{k}b}(n/\check{k}) \ln n/k}{\sqrt{\check{k}b}(n/\check{k})} \rightarrow 0. \quad (8)$$

Concerning $\widehat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \widehat{\rho}_k)$, i.e. if we use the same k sequence for the estimation of D and ρ , we have that if $k, n \rightarrow \infty$ such that $k/n \rightarrow 0$, $\sqrt{kb}(n/k)/\ln n/k \rightarrow \infty$, $\sqrt{kb}(n/k)\tilde{b}(n/k) \rightarrow \lambda_1$ and $\sqrt{kb}^2(n/k) \rightarrow \lambda_2$ then

$$\frac{\sqrt{kb}(n/k)}{\ln n/k} \left[\widehat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \widehat{\rho}_k) - D \right] \xrightarrow{\mathcal{D}} N \left(-D(\lambda_1\beta \tilde{c}_1(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho, \beta) + \lambda_2\tilde{c}_2(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho)), D^2\tilde{v}^2(\mathbb{K}^{(1,2)}, \omega_1, \omega_2, l, \gamma, \rho) \right).$$

Note that when D and ρ are estimated using the same sequence of k then the normalized estimator of D inherits the asymptotic behavior of the normalized ρ estimator. For a detailed discussion about k sequences satisfying the conditions of Theorem 2, and that can be used when estimating ρ we refer to Caeiro and Gomes (2006), Caeiro *et al.* (2009) and Gomes *et al.* (2009).

3 Numerical results

In this section we compare several estimators for D , both at the asymptotic level and through small sample simulations. We consider examples of our class of estimators obtained with the power and log weight functions, namely $\widehat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \rho) := \widehat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, 1, 2, \rho)$ with $\mathbb{P}^{(1)} := (\mathbb{P}_{\delta_1}, \mathbb{P}_{\delta_2}, \mathbb{P}_{\delta_2}, \mathbb{P}_{\delta_4})$ and $\mathbb{P}^{(2)} := (\mathbb{P}_{\delta_3}, \mathbb{P}_{\delta_2}, \mathbb{P}_{\delta_2}, \mathbb{P}_{\delta_4})$, and $\widehat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \rho) := \widehat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, 1, 2, \rho)$ with $\mathbb{L}^{(1)} := (\mathbb{L}_{\tau_1}, \mathbb{L}_{\tau_2}, \mathbb{L}_{\tau_2}, \mathbb{L}_{\tau_4})$ and $\mathbb{L}^{(2)} := (\mathbb{L}_{\tau_3}, \mathbb{L}_{\tau_2}, \mathbb{L}_{\tau_2}, \mathbb{L}_{\tau_4})$ where for both the power and the log kernel functions we set $\delta_1 = \tau_1 = 0.25$, $\delta_2 = \tau_2 = 0.5$, $\delta_3 = \tau_3 = 0.75$ and $\delta_4 = \tau_4 = 1$. The motivation for using these values for the tuning parameters stems from a small preliminary investigation of the asymptotic mean squared errors of the resulting estimators, in which they provided a reasonable performance for a wide range of models. These values are however in no way optimally selected, so further improvements over the presented results are still possible. The parameter α_1 is not fixed in advance, and will be used to fine tune the estimator for a specific distribution. When reasonable choices are made for the other parameters, this provides sufficient flexibility for practical applications.

The performance of our estimators will be compared with the following estimators from the recent extreme value literature:

- The maximum likelihood estimator for D from Gomes and Martins (2002):

$$\widehat{D}_{n,k}^{(GM)}(\rho) := \left(\frac{k}{n}\right)^\rho \frac{\left[\frac{1}{k} \sum_{j=1}^k \left(\frac{j}{k}\right)^{-\rho}\right] \left[\frac{1}{k} \sum_{j=1}^k Z_j\right] - \frac{1}{k} \sum_{j=1}^k \left(\frac{j}{k}\right)^{-\rho} Z_j}{\left[\frac{1}{k} \sum_{j=1}^k \left(\frac{j}{k}\right)^{-\rho}\right] \left[\frac{1}{k} \sum_{j=1}^k \left(\frac{j}{k}\right)^{-\rho} Z_j\right] - \frac{1}{k} \sum_{j=1}^k \left(\frac{j}{k}\right)^{-2\rho} Z_j}.$$

- The moment estimator introduced by Caeiro and Gomes (2006):

$$\widehat{D}_{n,k}^{(CG,\tau)}(\rho) := \begin{cases} -\frac{2(2-\rho)^2}{\tau\rho} \left(\frac{k}{n}\right)^\rho \frac{\left[\left(M_n^{(1)}(k)\right)^\tau - \left(M_n^{(2)}(k)/2\right)^{\tau/2}\right]^2}{\left(M_n^{(2)}(k)/2\right)^\tau - \left(M_n^{(4)}(k)/24\right)^{\tau/2}}, & \text{if } \tau \neq 0, \\ -\frac{2(2-\rho)^2}{\rho} \left(\frac{k}{n}\right)^\rho \frac{\left[\ln(M_n^{(1)}(k)) - \frac{1}{2}\ln(M_n^{(2)}(k)/2)\right]^2}{\ln(M_n^{(2)}(k)/2) - \frac{1}{2}\ln(M_n^{(4)}(k)/24)}, & \text{if } \tau = 0, \end{cases}$$

where

$$M_n^{(\alpha)}(k) := \frac{1}{k} \sum_{j=1}^k (\ln X_{n-j+1,n} - \ln X_{n-k,n})^\alpha, \quad \alpha > 0.$$

The latter two estimators have an asymptotic behavior that is in line with that stated in Theorem 2 for $\widehat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho)$, namely under $(\widetilde{\mathcal{R}})$, for $k, n \rightarrow \infty$, $k/n \rightarrow 0$, $\sqrt{k}b(n/k) \rightarrow \infty$, $\sqrt{k}b(n/k)\tilde{b}(n/k) \rightarrow \lambda_1$ and $\sqrt{k}b^2(n/k) \rightarrow \lambda_2$,

$$\sqrt{k}b(n/k)(\widehat{D}_{n,k}^{(\bullet)}(\rho) - D) \xrightarrow{\mathcal{D}} N(\lambda_1 u_D^{(\bullet)} + \lambda_2 v_D^{(\bullet)}, (w_D^{(\bullet)})^2), \quad (9)$$

where $\bullet$ is either GM or CG, τ ; see Gomes and Martins (2002) and Caeiro and Gomes (2006) for further details on $u_D^{(\bullet)}$, $v_D^{(\bullet)}$ and $w_D^{(\bullet)}$.

We start with an evaluation of the estimators at the asymptotic level. In Figure 1 we show the components u_D and v_D of the asymptotic bias, and the asymptotic standard deviation w_D of the estimators under study as a function of ρ for Pareto-type distributions with $\rho = \beta$ and $\gamma = D = 1$. Such parameter setting is satisfied by e.g. the Burr distribution (see Table 1). The estimator $\widehat{D}_{n,k}^{(GM)}(\rho)$ is a maximum likelihood estimator and hence has minimal asymptotic variance though through an appropriate selection of the tuning parameter α_1 the estimators $\widehat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \rho)$ and $\widehat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \rho)$ do not lose much. We set $\alpha_1 = 3.75$ and $\alpha_1 = -3.1$ for the estimators based on the power and the log kernel functions, respectively. These values give a good asymptotic efficiency relative to $\widehat{D}_{n,k}^{(CG,-1.2)}(\rho)$ for a wide range of ρ values; see the discussion about the asymptotic relative efficiency given below. The value $\tau = -1.2$ is taken from Caeiro and Gomes (2006) where it showed in the simulation experiment a good performance for a wide range of Burr models. Compared to $\widehat{D}_{n,k}^{(GM)}(\rho)$, $\widehat{D}_{n,k}(\mathbb{P}^{(1,2)}, 3.75, \rho)$ and $\widehat{D}_{n,k}(\mathbb{L}^{(1,2)}, -3.1, \rho)$, the estimator $\widehat{D}_{n,k}^{(CG,-1.2)}(\rho)$ is characterized by a high asymptotic standard deviation. For what concerns the asymptotic bias, there are two components, u_D and v_D , and therefore the performance of the estimators with respect to bias depends on the way these two components are combined in the overall bias. From (7) we obtain the following first order approximation to the asymptotic bias:

$$Abias(\widehat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho)) = \tilde{b}(n/k)u_D(\mathbb{K}^{(1,2)}, \theta, \rho, \beta) + b(n/k)v_D(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho).$$

Similar expressions are obtained for $\widehat{D}_{n,k}^{(GM)}(\rho)$ and $\widehat{D}_{n,k}^{(CG,\tau)}(\rho)$ based on (9). This is illustrated in Figure 2 with the Burr(1,1,1) distribution, having $\gamma = D = \xi = 1$ and $\rho = \beta = -1$ (see also Table

1). For distributions with $\rho = \beta$ the bias components of $\hat{D}_{n,k}^{(GM)}(\rho)$ satisfy $v_D^{(GM)} = -u_D^{(GM)}/\gamma$, which implies that the asymptotic bias is zero for instance for the Burr and GPD distributions. Consequently, for the class of Burr distributions the estimator $\hat{D}_{n,k}^{(GM)}(\rho)$ is asymptotically the optimal one. From Figure 2 it is clear that for an appropriate selection of the tuning parameter α_1 the asymptotic bias and variance of our estimators can be kept small, so that asymptotically they perform almost as good as the maximum likelihood estimator. An alternative way to compare the asymptotic performance of $\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \rho)$, where $\mathbb{K}^{(1,2)}$ is either $\mathbb{P}^{(1,2)}$ or $\mathbb{L}^{(1,2)}$, and $\hat{D}_{n,k}^{(CG,\tau)}(\rho)$, for distributions with $\rho = \beta$, is by considering their asymptotic relative efficiency (AREFF):

$$\begin{aligned} AREFF(\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \rho), \hat{D}_{n,k}^{(CG,\tau)}(\rho)) &:= \sqrt{\frac{\min_k AMSE(\hat{D}_{n,k}^{(CG,\tau)}(\rho))}{\min_k AMSE(\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \rho))}} \\ &= \sqrt{\left(\frac{(w_D^{(CG,\tau)})^2}{w_D^2(\mathbb{K}^{(1,2)}, 2, \rho)} \right)^{-2\rho/(1-4\rho)} \left(\frac{(B^{(CG,\tau)})^2}{B^2(\mathbb{K}^{(1,2)}, \alpha_1, \rho)} \right)^{(1-2\rho)/(1-4\rho)}}, \end{aligned}$$

where

$$\begin{aligned} B(\mathbb{K}^{(1,2)}, \alpha_1, \rho) &:= \xi u_D(\mathbb{K}^{(1,2)}, 2, \rho, \rho) + \gamma Dv_D(\mathbb{K}^{(1,2)}, \alpha_1, 1, 2, \rho), \\ B^{(CG,\tau)} &:= \xi u_D^{(CG,\tau)} + \gamma Dv_D^{(CG,\tau)}. \end{aligned}$$

In Figure 3 we show $AREFF(\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 3.75, \rho), \hat{D}_{n,k}^{(CG,-1.2)}(\rho))$ and $AREFF(\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -3.1, \rho), \hat{D}_{n,k}^{(CG,-1.2)}(\rho))$ as a function of ρ for the Burr(ζ, λ, δ) family. Our estimators show clearly a very attractive behavior over a wide range of ρ values. As mentioned above, $\hat{D}_{n,k}^{(GM)}(\rho)$ is asymptotically optimal for the Burr distributions, and therefore it is not included in the comparison based on AREFF.

In case of the $|T_\nu|$ distribution, where $|T_\nu|$ refers here to the absolute value of a Student t random variable with ν degrees of freedom, the asymptotic bias of $\hat{D}_{n,k}^{(GM)}(\rho)$ will not necessarily be equal to zero, and hence the estimators $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \rho)$, $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \rho)$ and $\hat{D}_{n,k}^{(CG,\tau)}(\rho)$ may asymptotically outperform the former. In Figure 4 we show the asymptotic bias, standard deviation and mean squared error of $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho)$, $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho)$, $\hat{D}_{n,k}^{(GM)}(\rho)$ and $\hat{D}_{n,k}^{(CG,-0.25)}(\rho)$ as a function of k in case of a sample of size $n = 5000$ from $|T_2|$. The tuning parameters were chosen in such a way that the asymptotic bias and standard deviation were both reasonably small, though the used values are not optimal. From Figure 4 it is clear that $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho)$, $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho)$ and $\hat{D}_{n,k}^{(CG,-0.25)}(\rho)$ are nearly asymptotically unbiased, and that $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho)$ and $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho)$ have an asymptotic standard deviation that is close to that of the maximum likelihood estimator $\hat{D}_{n,k}^{(GM)}(\rho)$. The asymptotic bias of $\hat{D}_{n,k}^{(GM)}(\rho)$ is considerable and therefore $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho)$ and $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho)$ reach the smallest minimum AMSE of the estimators under consideration. The plot of $AREFF(\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho), \hat{D}_{n,k}^{(CG,-0.25)}(\rho))$ and $AREFF(\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho), \hat{D}_{n,k}^{(CG,-0.25)}(\rho))$ versus ν given in Figure 5 illustrates also here the very attractive asymptotic behavior of

$\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho)$ and $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho)$ versus $\hat{D}_{n,k}^{(CG, -0.25)}(\rho)$ for a wide range of $|T_\nu|$ distributions. Compared to $\hat{D}_{n,k}^{(GM)}(\rho)$, the estimators $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho)$ and $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho)$ perform only better for $\nu \leq 6$, corresponding to $\gamma \geq 1/6$ and $\rho \leq -1/3$.

Note: The expressions for the bias components suggest that in case $\rho = \beta$ we can make the asymptotic bias of the estimators $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \rho)$, $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \rho)$ and $\hat{D}_{n,k}^{(CG, \tau)}$ zero by an appropriate choice of α_1 and τ , respectively, and hence for an asymptotic evaluation of the estimators it suffices in principle to compare the asymptotic standard deviations. However, the bias-eliminating values for the tuning parameters are distribution dependent, and are in practical applications unknown, so therefore it is relevant to have values for the tuning parameters that work well for a reasonable range of members of a particular family of distributions, which motivates the above analyses based on AREFF.

It is a well known fact that when estimating higher order parameters, the differences between the asymptotic and finite sample results may become considerable (Draisma, 2000, p 43–59, Gomes *et al.*, 2002). Therefore, we also investigate the properties of the four considered estimators with a small simulation study. The distributions considered in the simulation are

- the Burr(1,1,1) distribution: $\gamma = 1$, $\rho = -1$, $D = 1$;
- the Burr(1,0.5,2) distribution: $\gamma = 1$, $\rho = -2$, $D = 1$;
- the Fréchet(1) distribution: $\gamma = 1$, $\rho = -1$, $D = 0.5$;
- the $|T_1|$ distribution: $\gamma = 1$, $\rho = -2$, $D = 1.645$;
- the $|T_2|$ distribution: $\gamma = 0.5$, $\rho = -1$, $D = 1.5$;

we refer to Table 1 for further details. For each of the considered distributions we generated 1000 data sets of size $n = 5000$, and computed the four estimators under study for $k = 10$ to $k = 4990$ in steps of 10. The estimators are implemented with the power kernel ρ estimator of Goegebeur *et al.* (2010), with $\check{k} = \lfloor n^{0.95} \rfloor$ in case $\rho = -1$ and $\check{k} = \lfloor n^{0.975} \rfloor$ in case $\rho = -2$. Note that, in case one would use the AMSE optimal sequence of k for the estimation of D , given by $k = O(n^{-4\rho/(1-4\rho)})$ (according to (7)), these choices for $\check{k}$ would satisfy condition (8). We refer to Caeiro *et al.* (2009) for a more detailed discussion about the choice of $\check{k}$. The results are reported in Figures 6 to 9. The considered values of α_1 for $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \hat{\rho}_{\check{k}})$ and $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \hat{\rho}_{\check{k}})$ were inspired by the above asymptotic evaluations. For what concerns $\hat{D}_{n,k}^{(CG, \tau)}(\rho)$ we set τ at the value that worked best in the simulations reported in Caeiro and Gomes (2006), i.e. -1.2 for Burr distributions and -3 for the Fréchet(1) distribution. For the $|T_1|$ and $|T_2|$ distributions we tried several values and report here only the results for $\tau = -0.25$, which worked best.

Although the simulation experiment is limited we can draw the following provisional conclusions:

- Consistent with the asymptotic evaluation we find that $\hat{D}_{n,k}^{(GM)}(\hat{\rho}_{\check{k}})$ shows the best performance in terms of minimal MSE on the Burr distributions, though the difference with $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 3.25, \hat{\rho}_{\check{k}})$ is small. For the Fréchet(1), $|T_1|$ and $|T_2|$ distributions, $\hat{D}_{n,k}^{(GM)}(\hat{\rho}_{\check{k}})$

has considerable bias, and is outperformed in terms of minimal MSE by $\widehat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$, $\widehat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$ and $\widehat{D}_{n,k}^{(CG,\tau)}(\widehat{\rho}_{\tilde{k}})$, for appropriate choices of the tuning parameters α_1 and τ . Despite its considerable bias on the Fréchet distribution, the estimator $\widehat{D}_{n,k}^{(GM)}(\widehat{\rho}_{\tilde{k}})$ performs reasonably well in terms of minimal MSE.

- Unlike the estimation of the second order rate parameter ρ , the performance of the estimators for D depends on the appropriate selection of the tuning parameter α_1 , an observation that was also made in Caeiro and Gomes (2006) concerning their tuning parameter τ . It is however rather easy to find values for the tuning parameters that give a good performance for a wide range of distributions within a particular family.
- The estimators $\widehat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$ and $\widehat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$ have sample paths that are more stable than those of $\widehat{D}_{n,k}^{(CG,\tau)}(\widehat{\rho}_{\tilde{k}})$, an observation that is in line with the smaller asymptotic variance of the former two found in the above asymptotic evaluation.
- The best (in terms of minimal MSE, and considered over the values of α_1) estimator $\widehat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$ seems to perform slightly better than the best $\widehat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$.
- Overall our estimators are competitive with $\widehat{D}_{n,k}^{(GM)}(\widehat{\rho}_{\tilde{k}})$ and $\widehat{D}_{n,k}^{(CG,\tau)}(\widehat{\rho}_{\tilde{k}})$, though none of the considered estimators performs uniformly best. For practical applications we therefore suggest to consider several estimators for the parameter D and to use the common stable part of their sample paths to infer about the true value of D . Further work is in progress on the appropriate selection of the tuning parameters, and on the evaluation of other kernel functions.

Finally we illustrate the use of the considered weighted estimators for D on bias corrected estimation of the extreme value index γ . For the Hill estimator (Hill, 1975), $H_{k,n} := T_{n,k}(K)$, with $K(u) = 1_{(0,1)}(u)$, we have under the second order condition $(\mathcal{R})$, and for $k, n \rightarrow \infty$, $k/n \rightarrow 0$ and $\sqrt{k}b(n/k) \rightarrow \lambda$ that

$$\sqrt{k}(H_{k,n} - \gamma) \xrightarrow{\mathcal{D}} N\left(\frac{\lambda}{1-\rho}, \gamma^2\right), \quad (10)$$

from which we deduce the following bias-corrected estimator

$$H_{k,n}^{(BC)} := H_{k,n} \left(1 - \frac{D}{1-\rho} \left(\frac{n}{k}\right)^\rho\right).$$

This bias-corrected estimator for γ was first considered in Caeiro *et al.* (2005). We now examine the finite sample behavior of this estimator implemented with the power weighted estimator for ρ of Goegebeur *et al.* (2010), and the estimators $\widehat{D}_{n,\tilde{k}}(\mathbb{P}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$, $\widehat{D}_{n,\tilde{k}}(\mathbb{L}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$, $\widehat{D}_{n,\tilde{k}}^{(GM)}(\widehat{\rho}_{\tilde{k}})$ and $\widehat{D}_{n,\tilde{k}}^{(CG,\tau)}(\widehat{\rho}_{\tilde{k}})$ for D , in a small simulation experiment. In this the tuning parameters of the estimators for D are set as above. Note that, as in Caeiro and Gomes (2008) and Caeiro *et al.* (2009), we estimate D and ρ at the same level of $\tilde{k}$, which is as above taken as $\tilde{k} = \lfloor n^{0.95} \rfloor$ in case $\rho \geq -1$ and $\tilde{k} = \lfloor n^{0.975} \rfloor$ otherwise. We consider the distributions

Burr(1, 2, 0.5), having $\gamma = 1$ and $\rho = -0.5$, Burr(1, 1, 1), having $\gamma = 1$ and $\rho = -1$, and $|T_1|$, having $\gamma = 1$ and $\rho = -2$. For each of these distributions we simulate 1000 data sets of size $n = 1000$, and compute the estimators for $k = 5, 6, \dots, 999$. In Figure 10 we show the mean and the empirical mean squared error (MSE) of the Hill estimator and its bias-corrected versions, as a function of k . As is clear from Figure 10, the bias of the Hill estimator becomes more problematic as ρ gets closer to zero, an observation which is in line with the asymptotic result stated in (10). Also, the bias correction used in $H_{k,n}^{(BC)}$ is effective in that the estimators for γ show a longer stable portion around the true value, though the Burr distribution with $\rho = -0.5$ remains difficult. For what concerns the MSE it is clear that the use of estimator $\widehat{D}_{n,\tilde{k}}(\mathbb{L}^{(1,2)}, \alpha_1, \widehat{\rho}_{\tilde{k}})$ in $H_{k,n}^{(BC)}$ is highly competitive compared to $\widehat{D}_{n,\tilde{k}}^{(GM)}(\widehat{\rho}_{\tilde{k}})$, despite the asymptotic optimality of the latter on Burr distributions.

4 Conclusion

In this paper we introduced a very flexible class of estimators for the second order scale parameter, which has weighted sums of scaled log-spacings of successive order statistics as basic building blocks. Its consistency and asymptotic normality were established under appropriate conditions on the underlying distribution, the weight functions, and the sequences k and n . The asymptotic and finite sample evaluations indicated that the members from the class obtained for the power and log weight functions show an attractive behavior compared to the alternatives from the recent literature. Further work is in progress on, among others, the evaluation of other weight functions, the development of algorithms for the automatic determination of the tuning parameters, and for the selection of k .

5 Appendix

5.1 Proof of Theorem 1

Let $E_1, \dots, E_k$ denote independent unit exponential random variables. Under the conditions of Theorem 1, Goegebeur *et al.* (2010) derived the following asymptotic distributional representation

$$T_{n,k}(K) \stackrel{\mathcal{D}}{=} \gamma\mu(K) + \gamma\sigma(K) \frac{N_k(K)}{\sqrt{k}} (1 + o_{\mathbb{P}}(1)) + b(n/k)I(K, \rho)(1 + o_{\mathbb{P}}(1)),$$

with

$$N_k(K) := \sqrt{k} \frac{\frac{1}{k} \sum_{j=1}^k K\left(\frac{j}{k+1}\right) (E_j - 1)}{\sigma(K)},$$

which is asymptotically a standard normal random variable. Straightforward application of Taylor's theorem gives

$$\left(\frac{T_{n,k}(K)}{\mu(K)} \right)^{\alpha_1} \stackrel{\mathcal{D}}{=} \gamma^{\alpha_1} + \alpha_1 \gamma^{\alpha_1} \bar{\sigma}(K) \frac{N_k(K)}{\sqrt{k}} (1 + o_{\mathbb{P}}(1)) + b(n/k) \alpha_1 \gamma^{\alpha_1 - 1} \bar{I}(K, \rho)(1 + o_{\mathbb{P}}(1)),$$

and thus

$$\frac{\left(\frac{T_{n,k}(K_1)}{\mu(K_1)}\right)^{\alpha_1} - \left(\frac{T_{n,k}(K_2)}{\mu(K_2)}\right)^{\alpha_1}}{b(n/k)} \stackrel{\mathcal{D}}{=} \alpha_1 \gamma^{\alpha_1-1} \bar{I}^{(1)}(K_1, K_2, \rho)(1 + o_{\mathbb{P}}(1)) + \alpha_1 \gamma^{\alpha_1} \frac{N_k(K_1, K_2)}{\sqrt{k}b(n/k)}(1 + o_{\mathbb{P}}(1)), \quad (11)$$

where $N_k(K_1, K_2) := \bar{\sigma}(K_1)N_k(K_1) - \bar{\sigma}(K_2)N_k(K_2)$. By another application of Taylor's theorem we obtain

$$\begin{aligned} \left[\frac{\left(\frac{T_{n,k}(K_1)}{\mu(K_1)}\right)^{\alpha_1} - \left(\frac{T_{n,k}(K_2)}{\mu(K_2)}\right)^{\alpha_1}}{b(n/k)} \right]^{\theta} &\stackrel{\mathcal{D}}{=} \alpha_1^{\theta} \gamma^{\theta(\alpha_1-1)} [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta} (1 + o_{\mathbb{P}}(1)) \\ &\quad + \theta \alpha_1^{\theta} \gamma^{\theta(\alpha_1-1)+1} [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta-1} \frac{N_k(K_1, K_2)}{\sqrt{k}b(n/k)} (1 + o_{\mathbb{P}}(1)). \end{aligned} \quad (12)$$

By combining (11), (12) and using the fact that $b(x) = \gamma D x^{\rho}$ we have, if $k, n \rightarrow \infty$ such that $k/n \rightarrow 0$ and $\sqrt{k}b(n/k) \rightarrow \infty$

$$\left(\frac{k}{n}\right)^{\rho(\theta-1)} \Psi_{n,k}(\mathbb{K}^{(1)}, \alpha_1, \theta) \xrightarrow{\mathbb{P}} \frac{(\alpha_1 D)^{\theta-1} [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta}}{\theta \bar{I}^{(1)}(K_3, K_4, \rho)},$$

and similarly,

$$\left(\frac{k}{n}\right)^{\rho(\theta-2)} \Psi_{n,k}(\mathbb{K}^{(2)}, \alpha_2, \theta-1) \xrightarrow{\mathbb{P}} \frac{(\alpha_2 D)^{\theta-2} [\bar{I}^{(1)}(K_5, K_6, \rho)]^{\theta-1}}{(\theta-1) \bar{I}^{(1)}(K_7, K_8, \rho)}.$$

This then yields

$$\left(\frac{k}{n}\right)^{\rho} \Lambda_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta) \xrightarrow{\mathbb{P}} D \frac{(\theta-1) \alpha_1^{\theta-1} \bar{I}^{(1)}(K_7, K_8, \rho) [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta}}{\theta \alpha_2^{\theta-2} \bar{I}^{(1)}(K_3, K_4, \rho) [\bar{I}^{(1)}(K_5, K_6, \rho)]^{\theta-1}},$$

from which the consistency of $\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho)$ for D follows.

In case ρ is replaced by a consistent estimator, we apply Taylor's theorem and obtain

$$\begin{aligned} \hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \hat{\rho}_k) &= \\ \hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) &- \hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) \ln(n/k) (\hat{\rho}_k - \rho) (1 + o_{\mathbb{P}}(1)), \end{aligned} \quad (13)$$

so that we have consistency if $\ln(n/k)(\hat{\rho}_k - \rho) = o_{\mathbb{P}}(1)$. $\square$

5.2 Some lemmas

The first lemma gives the asymptotic expansion of $T_{n,k}(K)$ under $(\tilde{\mathcal{R}})$ and $(\tilde{\mathcal{K}})$. Let $Y_{1,n} \leq \dots \leq Y_{n,n}$ denote the order statistics of a random sample of size n from the unit Pareto distribution.

Lemma 1 Let $X_1, \dots, X_n$ be i.i.d. random variables according to a distribution satisfying $(\tilde{\mathcal{R}})$. If $(\tilde{\mathcal{K}})$ holds, then for $k, n \rightarrow \infty$ such that $k/n \rightarrow 0$ we have

$$\begin{aligned} T_{n,k}(K) &\stackrel{\mathcal{D}}{=} \gamma\mu(K) + \gamma\sigma(K)\frac{N_k(K)}{\sqrt{k}}(1 + o_{\mathbb{P}}(1)) + b(Y_{n-k,n})I(K, \rho) \\ &\quad + b(Y_{n-k,n})\tilde{b}(Y_{n-k,n})I(K, \rho + \beta)(1 + o_{\mathbb{P}}(1)) + b(Y_{n-k,n})O_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right), \end{aligned}$$

where $N_k(K)$ is an asymptotic standard normally distributed random variable.

Proof:

The proof of the lemma proceeds along the lines of argumentation used in Theorem 2 of Goegebeur *et al.* (2010), where the asymptotic expansion of $T_{n,k}(K)$ is derived under a slightly different version of condition $(\tilde{\mathcal{R}})$, and therefore we focus here only on the development of the terms that are different.

From the inverse probability integral transform we have that $X_{j,n} \stackrel{\mathcal{D}}{=} U(Y_{j,n})$, $j = 1, \dots, n$, where $Y_{1,n} \leq \dots \leq Y_{n,n}$ denote the order statistics of a random sample of size n from the unit Pareto distribution. Further $Y_{n-j+1,n}/Y_{n-k,n} \stackrel{\mathcal{D}}{=} Y_{k-j+1,k}$, $j = 1, \dots, k$, independently from $Y_{n-k,n}$. This enables us to write

$$\begin{aligned} \ln X_{n-j+1,n} - \ln X_{n-k,n} &\stackrel{\mathcal{D}}{=} \gamma \ln Y_{k-j+1,k} + \ln \frac{\ell_U(Y_{k-j+1,k}Y_{n-k,n})}{\ell_U(Y_{n-k,n})} \\ &= \gamma \ln Y_{k-j+1,k} + b(Y_{n-k,n})\frac{Y_{k-j+1,k}^\rho - 1}{\rho} \\ &\quad + b(Y_{n-k,n})\tilde{b}(Y_{n-k,n})\frac{Y_{k-j+1,k}^{\rho+\beta} - 1}{\rho + \beta} \\ &\quad + b(Y_{n-k,n})\tilde{b}(Y_{n-k,n})B_{n,k}(j), \end{aligned}$$

where

$$B_{n,k}(j) := \frac{\frac{\ln \ell_U(Y_{k-j+1,k}Y_{n-k,n}) - \ln \ell_U(Y_{n-k,n})}{b(Y_{n-k,n})} - \frac{Y_{k-j+1,k}^\rho - 1}{\rho}}{\tilde{b}(Y_{n-k,n})} - \frac{Y_{k-j+1,k}^{\rho+\beta} - 1}{\rho + \beta},$$

$j = 1, \dots, k$.

Using Rényi's representation (Rényi, 1953) for exponential order statistics we obtain

$$\begin{aligned}
T_{n,k}(K) &\stackrel{\mathcal{D}}{=} \gamma \frac{1}{k} \sum_{j=1}^k K \left(\frac{j}{k+1} \right) E_j + b(Y_{n-k,n}) \frac{1}{k} \sum_{j=1}^k K \left(\frac{j}{k+1} \right) j \frac{Y_{k-j+1,k}^\rho - Y_{k-j,k}^\rho}{\rho} \\
&\quad + b(Y_{n-k,n}) \tilde{b}(Y_{n-k,n}) \frac{1}{k} \sum_{j=1}^k K \left(\frac{j}{k+1} \right) j \frac{Y_{k-j+1,k}^{\rho+\beta} - Y_{k-j,k}^{\rho+\beta}}{\rho + \beta} \\
&\quad + b(Y_{n-k,n}) \tilde{b}(Y_{n-k,n}) \frac{1}{k} \sum_{j=1}^k K \left(\frac{j}{k+1} \right) \eta_j \\
&=: \tilde{T}_{n,k}^{(1)} + \tilde{T}_{n,k}^{(2)} + \tilde{T}_{n,k}^{(3)} + \tilde{T}_{n,k}^{(4)},
\end{aligned}$$

where $E_1, \dots, E_k$ are independent unit exponential random variables, and $\eta_j := j(B_{n,k}(j) - B_{n,k}(j+1))$, $j = 1, \dots, k$, $B_{n,k}(k+1) := 0$. The terms $\tilde{T}_{n,k}^{(1)}$ and $\tilde{T}_{n,k}^{(2)}$ are identical to those in Goegebeur *et al.* (2010) where it was shown that

$$\begin{aligned}
\tilde{T}_{n,k}^{(1)} &= \gamma \mu(K) + \gamma \sigma(K) \frac{N_k(K)}{\sqrt{k}} (1 + o_{\mathbb{P}}(1)), \\
\tilde{T}_{n,k}^{(2)} &= b(Y_{n-k,n}) I(K, \rho) + b(Y_{n-k,n}) O_{\mathbb{P}}(1/\sqrt{k}).
\end{aligned}$$

Note that the weighted average in $\tilde{T}_{n,k}^{(3)}$ is of the same form as that appearing in $\tilde{T}_{n,k}^{(2)}$, and hence

$$\tilde{T}_{n,k}^{(3)} = b(Y_{n-k,n}) \tilde{b}(Y_{n-k,n}) I(K, \rho + \beta) (1 + o_{\mathbb{P}}(1)).$$

Finally consider $\tilde{T}_{n,k}^{(4)}$. From $(\tilde{K})$ (i') we have

$$\begin{aligned}
\frac{1}{k} \sum_{j=1}^k K \left(\frac{j}{k+1} \right) \eta_j &= \frac{k+1}{k} \sum_{j=1}^k \frac{\eta_j}{j} \sum_{i=1}^j \int_{(i-1)/(k+1)}^{i/(k+1)} u(v) dv \\
&= \frac{k+1}{k} \sum_{i=1}^k \int_{(i-1)/(k+1)}^{i/(k+1)} u(v) dv \sum_{j=i}^k \frac{\eta_j}{j} \\
&= \frac{k+1}{k} \sum_{i=1}^k B_{n,k}(i) \int_{(i-1)/(k+1)}^{i/(k+1)} u(v) dv
\end{aligned}$$

and

$$\left| \frac{1}{k} \sum_{j=1}^k K \left(\frac{j}{k+1} \right) \eta_j \right| \leq \frac{1}{k} \sum_{i=1}^k f \left(\frac{i}{k+1} \right) |B_{n,k}(i)|.$$

Concerning $B_{n,k}(i)$ we use Drees (1998), from which it follows that for every $\varepsilon > 0$ there exists n_0 such that for any $n \geq n_0$, with arbitrary large probability, for $i = 1, \dots, k$,

$$\begin{aligned}
|B_{n,k}(i)| &\leq \varepsilon \max \left\{ Y_{k-i+1,k}^{\rho+\beta+\delta}, Y_{k-i+1,k}^{\rho+\beta-\delta} \right\} \\
&= \varepsilon Y_{k-i+1,k}^{\rho+\beta+\delta} \\
&\leq \varepsilon,
\end{aligned}$$

provided $0 < \delta < -\rho - \beta$, and thus

$$\sup_{i \in \{1, \dots, k\}} |B_{n,k}(i)| = o_{\mathbb{P}}(1),$$

leading to

$$\left| \frac{1}{k} \sum_{j=1}^k K\left(\frac{j}{k+1}\right) \eta_j \right| \leq o_{\mathbb{P}}(1) \frac{1}{k} \sum_{i=1}^k f\left(\frac{i}{k+1}\right),$$

which, by assumption $(\tilde{\mathcal{K}})(i')$ is $o_{\mathbb{P}}(1)$. Thus $\tilde{T}_{n,k}^{(4)} = b(Y_{n-k,n})\tilde{b}(Y_{n-k,n})o_{\mathbb{P}}(1)$.

The combination of the above results establishes Lemma 1. $\square$

Lemma 2 *Let $Y_{1,n} \leq \dots \leq Y_{n,n}$ denote the order statistics from a random sample of size n from the unit Pareto distribution. If $\rho < 0$ and $k, n \rightarrow \infty$ such that $k/n \rightarrow 0$ and $\sqrt{k}(n/k)^\rho \rightarrow \infty$ we have that*

$$\left(\frac{Y_{n-k,n}}{n/k}\right)^\rho \stackrel{\mathcal{D}}{=} 1 + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}(n/k)^\rho}\right).$$

Proof:

Let $E_{1,n} \leq \dots \leq E_{n,n}$ denote the order statistics of a random sample of size n from the unit exponential distribution. We have

$$\begin{aligned} Y_{n-k,n}^\rho &\stackrel{\mathcal{D}}{=} e^{\rho E_{n-k,n}} \\ &= e^{\rho \ln n/k} + \rho e^{\rho Q_{n,k}}(E_{n-k,n} - \ln n/k), \end{aligned}$$

where $Q_{n,k}$ is a random value between $E_{n-k,n}$ and $\ln n/k$. From Lemma 1 of Girard (2004) we have that $E_{n-k,n} - \ln n/k = O_{\mathbb{P}}(1/\sqrt{k})$ and hence $e^{\rho Q_{n,k}} = o_{\mathbb{P}}(1)$. This then gives

$$Y_{n-k,n}^\rho \stackrel{\mathcal{D}}{=} \left(\frac{n}{k}\right)^\rho + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right),$$

from which the result follows. $\square$

Lemma 3 *The kernel function*

$$K(u; \tau, \delta) = (-\ln u)^\tau u^\delta, \quad u \in (0, 1); \tau, \delta \geq 0,$$

satisfies conditions $(\mathcal{K})$ and $(\tilde{\mathcal{K}})$.

Proof:

For part (i) of $(\mathcal{K})$ we have that

$$u(t) = t^\delta [(1 + \delta)(-\ln t)^\tau - \tau(-\ln t)^{\tau-1}],$$

and hence

$$\left| (k+1) \int_{(j-1)/(k+1)}^{j/(k+1)} u(t) dt \right| \leq (k+1) \left[(1 + \delta) \int_{(j-1)/(k+1)}^{j/(k+1)} (-\ln t)^\tau dt + \tau \int_{(j-1)/(k+1)}^{j/(k+1)} (-\ln t)^{\tau-1} dt \right].$$

We distinguish two cases.

Case 1: $\tau \geq 1$. We have

$$\left| (k+1) \int_{(j-1)/(k+1)}^{j/(k+1)} u(t) dt \right| \leq \frac{k+1}{j} \left[(1 + \delta) \int_0^{j/(k+1)} (-\ln t)^\tau dt + \tau \int_0^{j/(k+1)} (-\ln t)^{\tau-1} dt \right],$$

and thus

$$f(x) = \frac{1 + \delta}{x} \int_0^x (-\ln t)^\tau dt + \frac{\tau}{x} \int_0^x (-\ln t)^{\tau-1} dt$$

which satisfies $\int_0^1 \ln^+(1/x) f(x) dx < \infty$.

Case 2: $\tau < 1$. In this case

$$\left| (k+1) \int_{(j-1)/(k+1)}^{j/(k+1)} u(t) dt \right| \leq (1 + \delta) \frac{k+1}{j} \int_0^{j/(k+1)} (-\ln t)^\tau dt + \tau \left(-\ln \frac{j}{k+1} \right)^{\tau-1},$$

so

$$f(x) = \frac{1 + \delta}{x} \int_0^x (-\ln t)^\tau dt + \tau (-\ln x)^{\tau-1},$$

also satisfying $\int_0^1 \ln^+(1/x) f(x) dx < \infty$.

Condition $(\mathcal{K})(ii)$ is trivially satisfied as $\int_0^1 K^2(u; \tau, \delta) du = \Gamma(1 + 2\tau)/(1 + 2\delta)^{1+2\tau} < \infty$.

Concerning $(\mathcal{K})(iii)$ observe that

$$\begin{aligned} \left| \frac{1}{k} \sum_{j=1}^k K\left(\frac{j}{k+1}; \tau, \delta\right) - \int_0^1 K(u; \tau, \delta) du \right| &\leq \left| \frac{1}{k+1} \sum_{j=1}^k K\left(\frac{j}{k+1}; \tau, \delta\right) - \int_{1/(k+1)}^1 K(u; \tau, \delta) du \right| \\ &\quad + \int_0^{1/(k+1)} K(u; \tau, \delta) du + O\left(\frac{1}{k}\right) \\ &= \left| \frac{1}{k+1} \sum_{j=1}^k \left[K\left(\frac{j}{k+1}; \tau, \delta\right) - K(\tilde{u}_j; \tau, \delta) \right] \right| \\ &\quad + O\left(\frac{(\ln k)^\tau}{k}\right), \end{aligned}$$

where $\tilde{u}_j$ is a value between $j/(k+1)$ and $(j+1)/(k+1)$. Thus

$$\begin{aligned} \left| \frac{1}{k+1} \sum_{j=1}^k \left[K\left(\frac{j}{k+1}; \tau, \delta\right) - K(\tilde{u}_j; \tau, \delta) \right] \right| &\leq \frac{1}{k+1} \sum_{j=1}^k \left(\frac{j}{k+1} \right)^\delta \left[\left(-\ln \frac{j}{k+1} \right)^\tau - (-\ln \tilde{u}_j)^\tau \right] \\ &\quad + \frac{1}{k+1} \sum_{j=1}^k (-\ln \tilde{u}_j)^\tau \left[\tilde{u}_j^\delta - \left(\frac{j}{k+1} \right)^\delta \right] \\ &=: L_1 + L_2. \end{aligned}$$

For L_1 and L_2 one easily obtains the following

$$\begin{aligned} L_1 &\leq \frac{1}{k+1} \sum_{j=1}^k \left[\left(-\ln \frac{j}{k+1} \right)^\tau - \left(-\ln \frac{j+1}{k+1} \right)^\tau \right] \\ &= \frac{(\ln(k+1))^\tau}{k+1} \\ L_2 &\leq \frac{(\ln(k+1))^\tau}{k+1} \sum_{j=1}^k \left[\left(\frac{j+1}{k+1} \right)^\delta - \left(\frac{j}{k+1} \right)^\delta \right] \\ &= O\left(\frac{(\ln k)^\tau}{k} \right). \end{aligned}$$

Combining the above we have that

$$\frac{1}{k} \sum_{j=1}^k K\left(\frac{j}{k+1}; \tau, \delta\right) = \int_0^1 K(u; \tau, \delta) du + O\left(\frac{(\ln k)^\tau}{k} \right),$$

and hence $(\mathcal{K})(iii)$ is satisfied.

As for $(\mathcal{K})(iv)$ we have

$$\max_{i \in \{1, \dots, k\}} K\left(\frac{i}{k+1}; \tau, \delta\right) \leq (\ln(k+1))^\tau,$$

and $(\ln(k+1))^\tau = o(\sqrt{k})$.

We now turn to the verification of $(\tilde{K})$. Condition (i') follows from $(\mathcal{K})(i)$ and (v) can be obtained from $(\mathcal{K})(iii)$. As for (vi) we easily obtain

$$\int_0^1 (-\ln u)^\tau u^{\delta+|\rho|-1-\varepsilon} du = \frac{\Gamma(1+\tau)}{(\delta+|\rho|-\varepsilon)^{1+\tau}},$$

provided $\varepsilon < |\rho| + \delta$. □

5.3 Proof of Theorem 2

From Lemma 1 and Taylor's theorem we obtain easily

$$\begin{aligned}
& \frac{\left(\frac{T_{n,k}(K_1)}{\mu(K_1)}\right)^{\alpha_1} - \left(\frac{T_{n,k}(K_2)}{\mu(K_2)}\right)^{\alpha_1}}{b(Y_{n-k,n})} \stackrel{\mathcal{D}}{=} \alpha_1 \gamma^{\alpha_1-1} \bar{I}^{(1)}(K_1, K_2, \rho) \\
& + \alpha_1 \gamma^{\alpha_1} \frac{N_k(K_1, K_2)}{\sqrt{k}b(Y_{n-k,n})} (1 + o_{\mathbb{P}}(1)) + \tilde{b}(Y_{n-k,n}) \alpha_1 \gamma^{\alpha_1-1} \bar{I}^{(1)}(K_1, K_2, \rho + \beta) (1 + o_{\mathbb{P}}(1)) \\
& + b(Y_{n-k,n}) \frac{\alpha_1(\alpha_1 - 1)}{2} \gamma^{\alpha_1-2} \bar{I}^{(2)}(K_1, K_2, \rho) (1 + o_{\mathbb{P}}(1)), \tag{14}
\end{aligned}$$

and

$$\begin{aligned}
& \left[\frac{\left(\frac{T_{n,k}(K_1)}{\mu(K_1)}\right)^{\alpha_1} - \left(\frac{T_{n,k}(K_2)}{\mu(K_2)}\right)^{\alpha_1}}{b(Y_{n-k,n})} \right]^{\theta} \stackrel{\mathcal{D}}{=} \alpha_1^{\theta} \gamma^{\theta(\alpha_1-1)} [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta} \\
& + \frac{1}{\sqrt{k}b(Y_{n-k,n})} \theta \alpha_1^{\theta} \gamma^{\theta(\alpha_1-1)+1} [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta-1} N_k(K_1, K_2) (1 + o_{\mathbb{P}}(1)) \\
& + \tilde{b}(Y_{n-k,n}) \theta \alpha_1^{\theta} \gamma^{\theta(\alpha_1-1)} [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta-1} \bar{I}^{(1)}(K_1, K_2, \rho + \beta) (1 + o_{\mathbb{P}}(1)) \\
& + b(Y_{n-k,n}) \theta \frac{\alpha_1^{\theta}(\alpha_1 - 1)}{2} \gamma^{\theta(\alpha_1-1)-1} [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta-1} \bar{I}^{(2)}(K_1, K_2, \rho) (1 + o_{\mathbb{P}}(1)).
\end{aligned}$$

From (14) we also derive

$$\begin{aligned}
& \frac{\left(\frac{T_{n,k}(K_3)}{\mu(K_3)}\right)^{\alpha_1 \theta} - \left(\frac{T_{n,k}(K_4)}{\mu(K_4)}\right)^{\alpha_1 \theta}}{b(Y_{n-k,n})} \stackrel{\mathcal{D}}{=} \alpha_1 \theta \gamma^{\alpha_1 \theta - 1} \bar{I}^{(1)}(K_3, K_4, \rho) \\
& + \alpha_1 \theta \gamma^{\alpha_1 \theta} \frac{N_k(K_3, K_4)}{\sqrt{k}b(Y_{n-k,n})} (1 + o_{\mathbb{P}}(1)) + \tilde{b}(Y_{n-k,n}) \alpha_1 \theta \gamma^{\alpha_1 \theta - 1} \bar{I}^{(1)}(K_3, K_4, \rho + \beta) (1 + o_{\mathbb{P}}(1)) \\
& + b(Y_{n-k,n}) \frac{\alpha_1 \theta (\alpha_1 \theta - 1)}{2} \gamma^{\alpha_1 \theta - 2} \bar{I}^{(2)}(K_3, K_4, \rho) (1 + o_{\mathbb{P}}(1)),
\end{aligned}$$

and hence

$$\begin{aligned}
& Y_{n-k,n}^{\rho(1-\theta)} \frac{\theta \bar{I}^{(1)}(K_3, K_4, \rho)}{\alpha_1^{\theta-1} [\bar{I}^{(1)}(K_1, K_2, \rho)]^{\theta}} \Psi_{n,k}(\mathbb{K}^{(1)}, \alpha_1, \theta) \stackrel{\mathcal{D}}{=} \\
& D^{\theta-1} + \frac{1}{\sqrt{k}b(Y_{n-k,n})} \gamma D^{\theta-1} \left[\frac{\theta N_k(K_1, K_2)}{\bar{I}^{(1)}(K_1, K_2, \rho)} - \frac{N_k(K_3, K_4)}{\bar{I}^{(1)}(K_3, K_4, \rho)} \right] (1 + o_{\mathbb{P}}(1)) \\
& + \tilde{b}(Y_{n-k,n}) D^{\theta-1} \left[\frac{\theta \bar{I}^{(1)}(K_1, K_2, \rho + \beta)}{\bar{I}^{(1)}(K_1, K_2, \rho)} - \frac{\bar{I}^{(1)}(K_3, K_4, \rho + \beta)}{\bar{I}^{(1)}(K_3, K_4, \rho)} \right] (1 + o_{\mathbb{P}}(1)) \\
& + b(Y_{n-k,n}) \frac{D^{\theta-1}}{2\gamma} \left[\frac{\theta(\alpha_1 - 1) \bar{I}^{(2)}(K_1, K_2, \rho)}{\bar{I}^{(1)}(K_1, K_2, \rho)} - \frac{(\alpha_1 \theta - 1) \bar{I}^{(2)}(K_3, K_4, \rho)}{\bar{I}^{(1)}(K_3, K_4, \rho)} \right] (1 + o_{\mathbb{P}}(1)).
\end{aligned}$$

Also

$$\begin{aligned}
& Y_{n-k,n}^{\rho(2-\theta)} \frac{(\theta-1)\bar{I}^{(1)}(K_7, K_8, \rho)}{\alpha_2^{\theta-2} [\bar{I}^{(1)}(K_5, K_6, \rho)]^{\theta-1}} \Psi_{n,k}(\mathbb{K}^{(2)}, \alpha_2, \theta-1) \stackrel{\mathcal{D}}{=} \\
& D^{\theta-2} + \frac{1}{\sqrt{k}b(Y_{n-k,n})} \gamma D^{\theta-2} \left[\frac{(\theta-1)N_k(K_5, K_6)}{\bar{I}^{(1)}(K_5, K_6, \rho)} - \frac{N_k(K_7, K_8)}{\bar{I}^{(1)}(K_7, K_8, \rho)} \right] (1 + o_{\mathbb{P}}(1)) \\
& + \tilde{b}(Y_{n-k,n}) D^{\theta-2} \left[\frac{(\theta-1)\bar{I}^{(1)}(K_5, K_6, \rho + \beta)}{\bar{I}^{(1)}(K_5, K_6, \rho)} - \frac{\bar{I}^{(1)}(K_7, K_8, \rho + \beta)}{\bar{I}^{(1)}(K_7, K_8, \rho)} \right] (1 + o_{\mathbb{P}}(1)) \\
& + b(Y_{n-k,n}) \frac{D^{\theta-2}}{2\gamma} \left[\frac{(\theta-1)(\alpha_2-1)\bar{I}^{(2)}(K_5, K_6, \rho)}{\bar{I}^{(1)}(K_5, K_6, \rho)} - \frac{(\alpha_2(\theta-1)-1)\bar{I}^{(2)}(K_7, K_8, \rho)}{\bar{I}^{(1)}(K_7, K_8, \rho)} \right] (1 + o_{\mathbb{P}}(1)),
\end{aligned}$$

so that by another application of Taylor's theorem we obtain

$$\begin{aligned}
& Y_{n-k,n}^{-\rho} c(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) \Lambda_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta) \stackrel{\mathcal{D}}{=} \\
& D + \frac{1}{\sqrt{k}b(Y_{n-k,n})} N_k(\mathbb{K}^{(1,2)}, \theta, \rho) (1 + o_{\mathbb{P}}(1)) + \tilde{b}(Y_{n-k,n}) u_D(\mathbb{K}^{(1,2)}, \theta, \rho, \beta) (1 + o_{\mathbb{P}}(1)) \\
& + b(Y_{n-k,n}) v_D(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho) (1 + o_{\mathbb{P}}(1)).
\end{aligned}$$

Finally, by using Lemma 2 we establish the asymptotic normality of the normalized $\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho)$.

The last part of Theorem 2, dealing with the replacement of ρ by an estimator in $\hat{D}_{n,k}(\mathbb{K}^{(1,2)}, \alpha_1, \alpha_2, \theta, \rho)$ follows easily from expansion (13). $\square$

Acknowledgement

The authors are very grateful to the referees for their careful reading of the paper and their comments that lead to significant improvements of the initial draft.

References

- [1] Beirlant, J., Dierckx, G., Goegebeur, Y., Matthys, G., 1999. Tail index estimation and an exponential regression model. *Extremes* 2, 177–200.
- [2] Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. *Regular Variation*. Cambridge University Press, Cambridge.
- [3] Caeiro, F., Gomes, M.I., 2006. A new class of estimators of a “scale” second order parameter. *Extremes* 9, 193–211.
- [4] Caeiro, F., Gomes, M.I., 2008. Minimum-variance reduced-bias tail index and high quantile estimation. *REVSTAT – Statistical Journal* 6, 1–20.
- [5] Caeiro, F., Gomes, M.I., Pestana, D., 2005. Direct reduction of bias of the classical Hill estimator. *REVSTAT – Statistical Journal* 3, 113–136.

- [6] Caeiro, F., Gomes, M.I., Rodrigues, L.H., 2009. Reduced-bias tail index estimators under a third order framework. *Communications in Statistics – Theory and Methods* 38, 1019–1040.
- [7] Ciuperca, G., Mercadier, C., 2010. Semi-parametric estimation for heavy tailed distributions. *Extremes* 13, 55–87.
- [8] Draisma, G., 2000. *Parametric and Semi-Parametric Methods in Extreme Value Theory*. Tinbergen Institute Research Series 239, Erasmus Universiteit Rotterdam.
- [9] Drees, H. 1998. On smooth statistical tail functionals. *Scandinavian Journal of Statistics* 25, 187–210.
- [10] Feuerverger, A., Hall, P., 1999. Estimating a tail exponent by modelling departure from a Pareto distribution. *Annals of Statistics* 27, 760–781.
- [11] Fisher, R.A., Tippett, L.H.C., 1928. Limiting form of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society* 24, 180–190.
- [12] Fraga Alves, M.I., Gomes, M.I., de Haan, L., 2003. A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica* 60, 193–213.
- [13] Fréchet, M., 1927. Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique* 6, 93–116.
- [14] Girard, S., 2004. A Hill type estimator of the Weibull tail-coefficient. *Communications in Statistics - Theory and Methods* 33, 205–234.
- [15] Gnedenko, B.V., 1943. Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics* 44, 423–453.
- [16] Goegebeur, Y., Beirlant, J., de Wet, T., 2008. Linking Pareto-tail kernel goodness-of-fit statistics with tail index at optimal threshold and second order estimation. *REVSTAT – Statistical Journal* 6, 51–69.
- [17] Goegebeur, Y., Beirlant, J., de Wet, T., 2010. Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference* 140, 2632–2652.
- [18] Goegebeur, Y., de Wet, T., 2011. Estimation of the third order parameter in extreme value statistics. To appear in *TEST*, DOI 10.1007/s11749-011-0246-2.
- [19] Gomes, M.I., de Haan, L., Peng, L., 2002. Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes* 5, 387–414.
- [20] Gomes, M.I., de Haan, L., Rodrigues, L.H., 2008. Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *Journal of the Royal Statistical Society Series B* 70, 31–52.

- [21] Gomes, M.I., Henriques-Rodrigues, L., Pereira, H., Pestana, D., 2010. Tail index and second order parameters' semi-parametric estimation based on the log-excesses. *Journal of Statistical Computation and Simulation* 80, 653–666.
- [22] Gomes, M.I., Martins, M.J., 2002. 'Asymptotically unbiased' estimators of the tail index based on external estimation of the second order parameter. *Extremes* 5, 5–31.
- [23] Gomes, M.I., Pestana, D., Caeiro, F., 2009. A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Statistics and Probability Letters* 79, 295–303.
- [24] de Haan, L., Ferreira, A., 2006. *Extreme Value Theory: An Introduction*. Springer.
- [25] Hall, P., Welsh, A.H., 1985. Adaptive estimates of parameters of regular variation. *Annals of Statistics* 13, 331–341.
- [26] Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Annals of Statistics* 3, 1163–1174.
- [27] Rényi, A., 1953. On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungaricae* 4, 191–231.

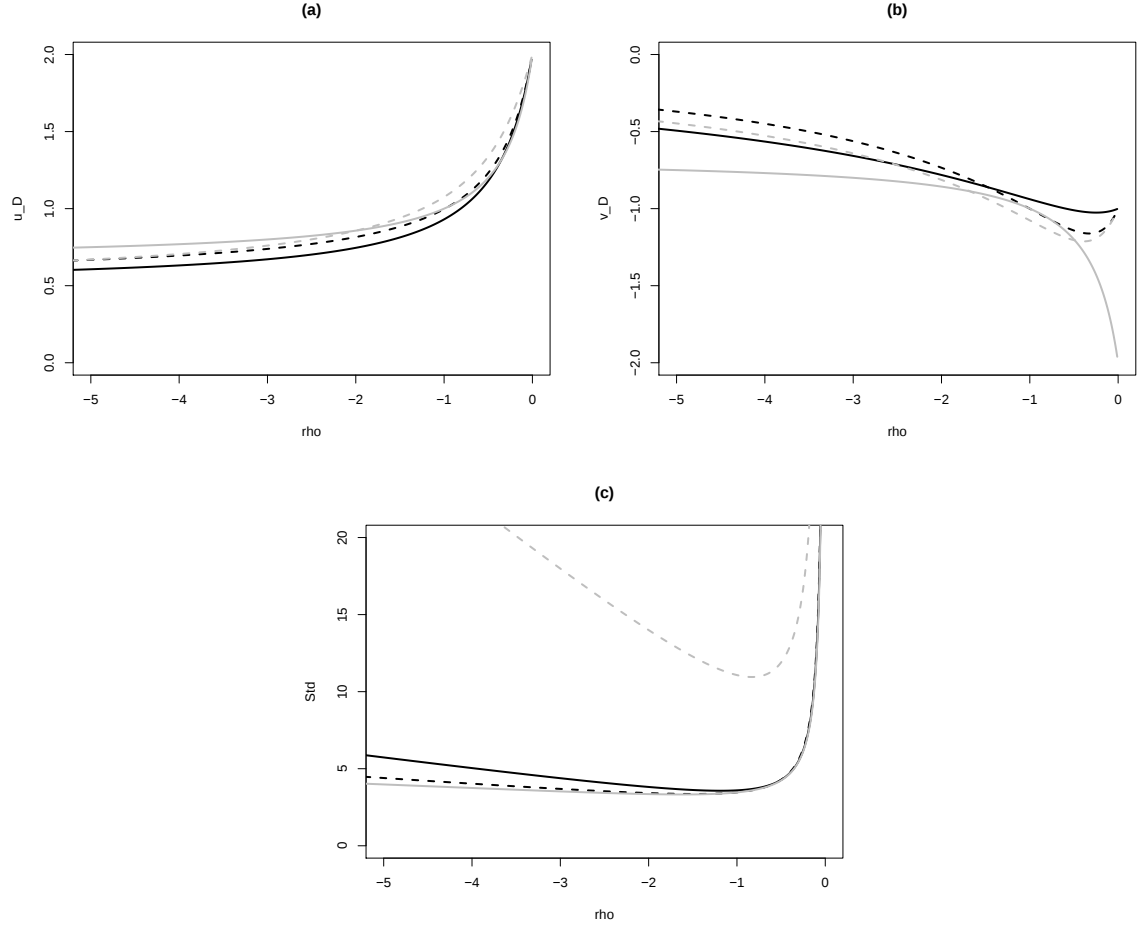


Figure 1: Pareto-type models with $\rho = \beta$ and $\gamma = D = 1$: (a) bias component u_D , (b) bias component v_D , and (c) asymptotic standard deviation as a function of ρ for $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 3.75, \rho)$ (black solid line), $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -3.1, \rho)$ (black dashed line), $\hat{D}_{n,k}^{(GM)}(\rho)$ (grey solid line) and $\hat{D}_{n,k}^{(CG, -1.2)}(\rho)$ (grey dashed line).

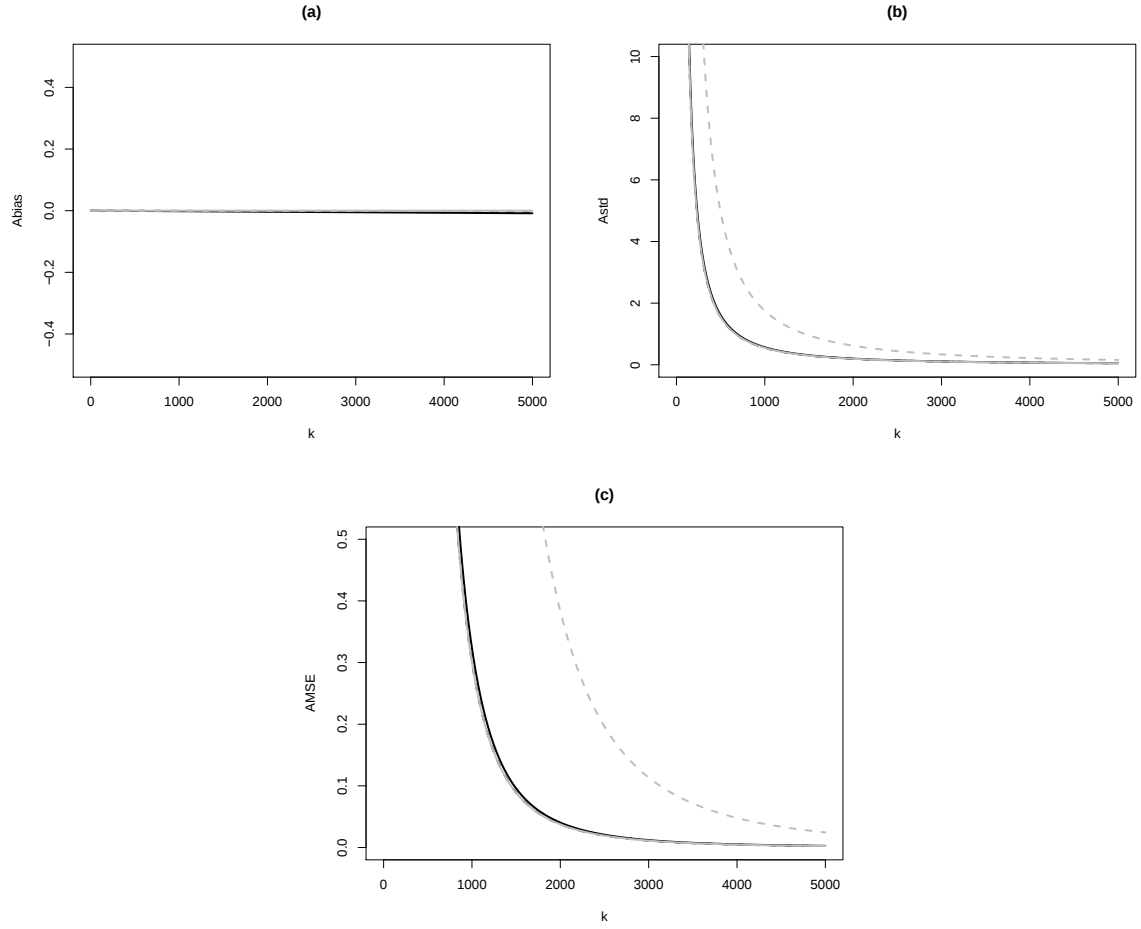


Figure 2: Burr(1,1,1) distribution: (a) asymptotic bias, (b) asymptotic standard deviation and (c) AMSE of $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 3.75, \rho)$ (black solid line), $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -3.1, \rho)$ (black dashed line), $\hat{D}_{n,k}^{(GM)}(\rho)$ (grey solid line) and $\hat{D}_{n,k}^{(CG, -1.2)}(\rho)$ (grey dashed line) as a function of k when $n = 5000$.

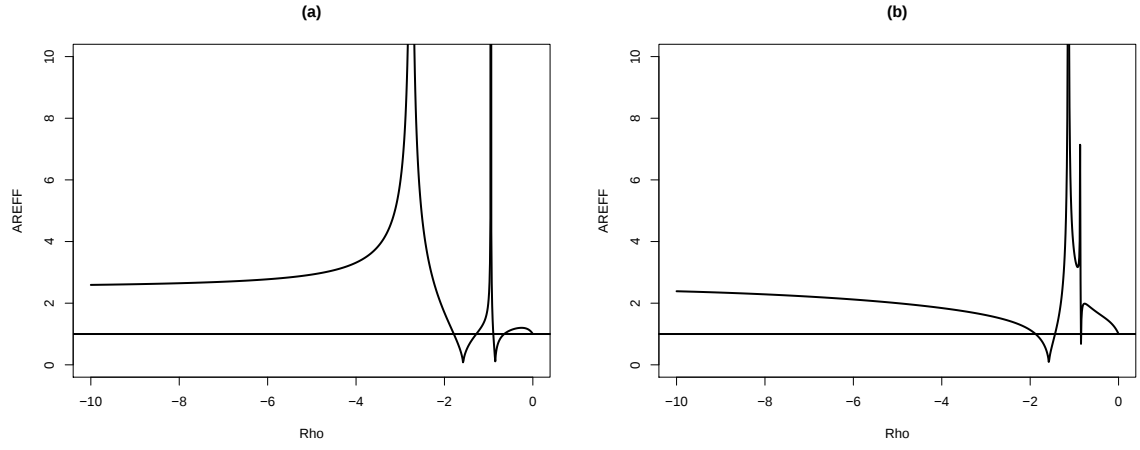


Figure 3: Burr(ζ, λ, δ) distribution: (a) $AREFF(\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 3.75, \rho), \hat{D}_{n,k}^{(CG, -1.2)}(\rho))$ and (b) $AREFF(\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -3.1, \rho), \hat{D}_{n,k}^{(CG, -1.2)}(\rho))$ as a function of ρ .

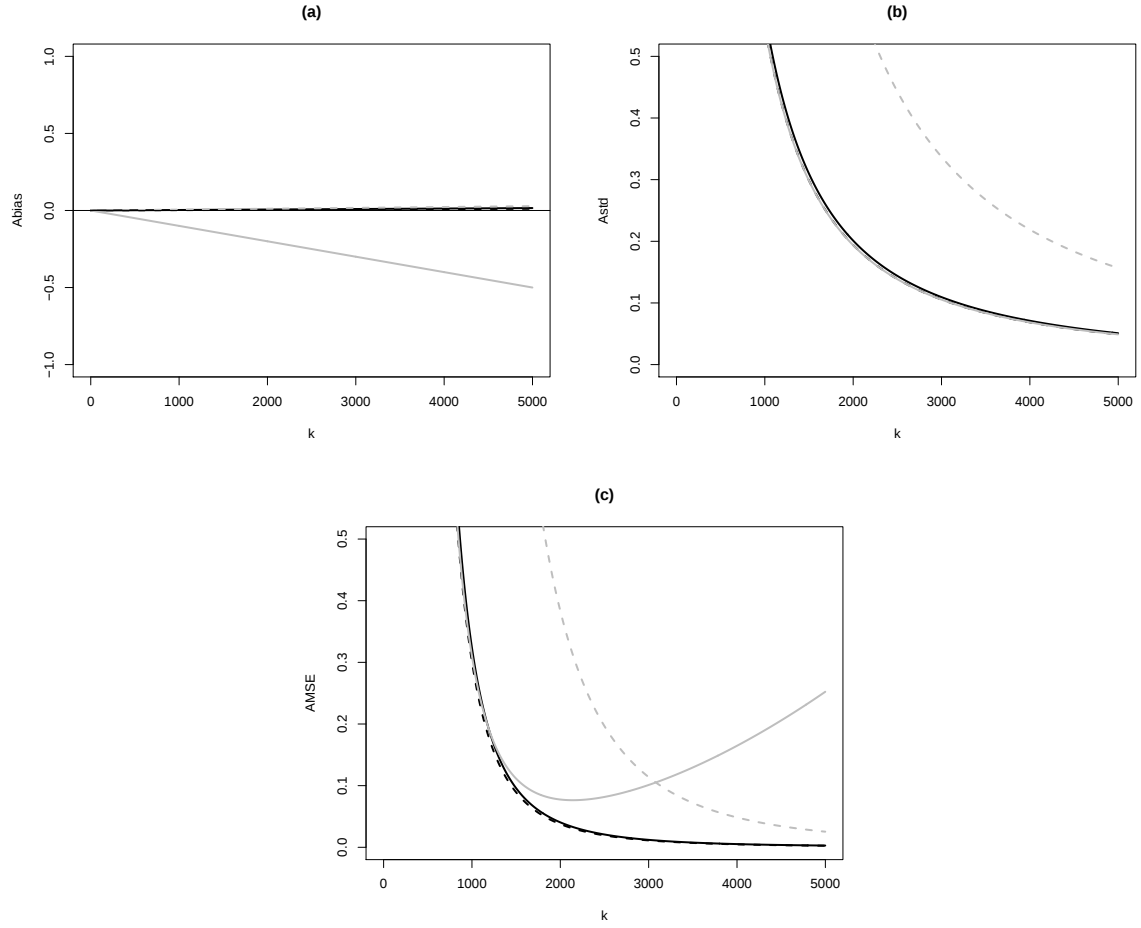


Figure 4: $|T_2|$ distribution: (a) asymptotic bias, (b) asymptotic standard deviation and (c) AMSE of $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho)$ (black solid line), $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho)$ (black dashed line), $\hat{D}_{n,k}^{(GM)}(\rho)$ (grey solid line) and $\hat{D}_{n,k}^{(CG, -0.25)}(\rho)$ (grey dashed line) as a function of k when $n = 5000$.

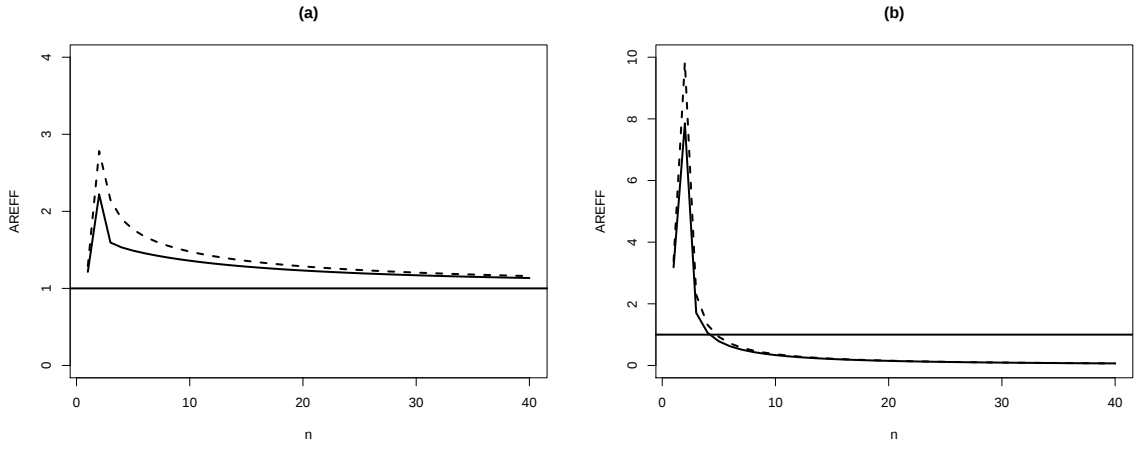


Figure 5: $|T_\nu|$ distribution: (a) $AREFF(\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho), \hat{D}_{n,k}^{(CG, -0.25)}(\rho))$ (solid line) and $AREFF(\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho), \hat{D}_{n,k}^{(CG, -0.25)}(\rho))$ (dashed line), and (b) $AREFF(\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, 1.75, \rho), \hat{D}_{n,k}^{(GM)}(\rho))$ (solid line) and $AREFF(\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, -1.75, \rho), \hat{D}_{n,k}^{(GM)}(\rho))$ (dashed line) as a function of ν .

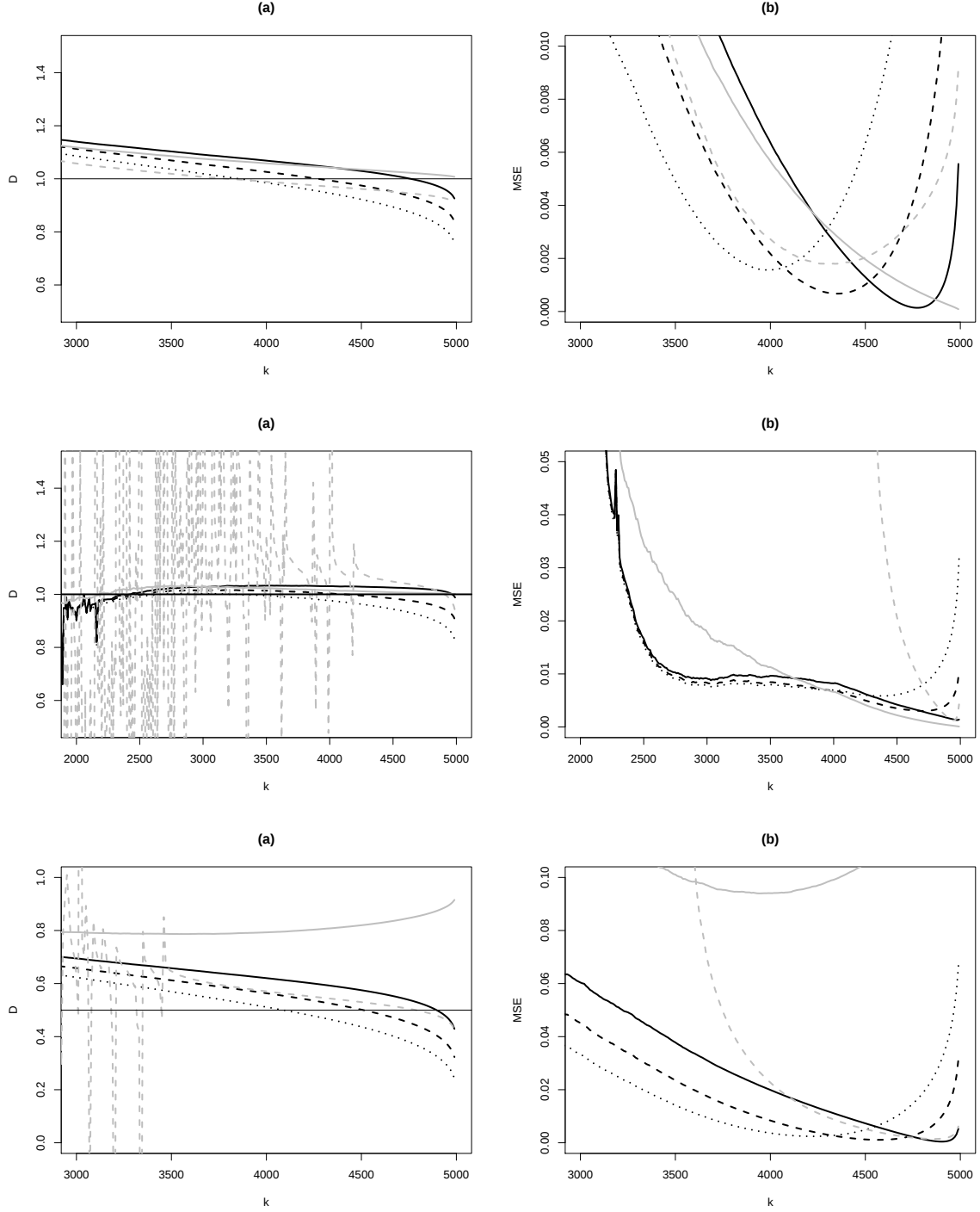


Figure 6: Sample mean (left) and empirical MSE (right) as a function of k when $n = 5000$. Burr(1,1,1) (top) and Burr(1,0.5,2) (middle) distributions: $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \hat{\rho}_k)$ with $\alpha_1 = 3.25$ (black solid line), 3.5 (black dashed line) and 3.75 (black dotted line), and $\hat{D}_{n,k}^{(CG,\tau)}(\hat{\rho}_k)$ with $\tau = -1.2$ (grey dashed line). Fréchet(1) distribution (bottom): $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \hat{\rho}_k)$ with $\alpha_1 = 6.0$ (black solid line), 7.0 (black dashed line) and 8.0 (black dotted line), and $\hat{D}_{n,k}^{(CG,\tau)}(\hat{\rho}_k)$ with $\tau = -3$ (grey dashed line). The grey solid line represents $\hat{D}_{n,k}^{(GM)}(\hat{\rho}_k)$.

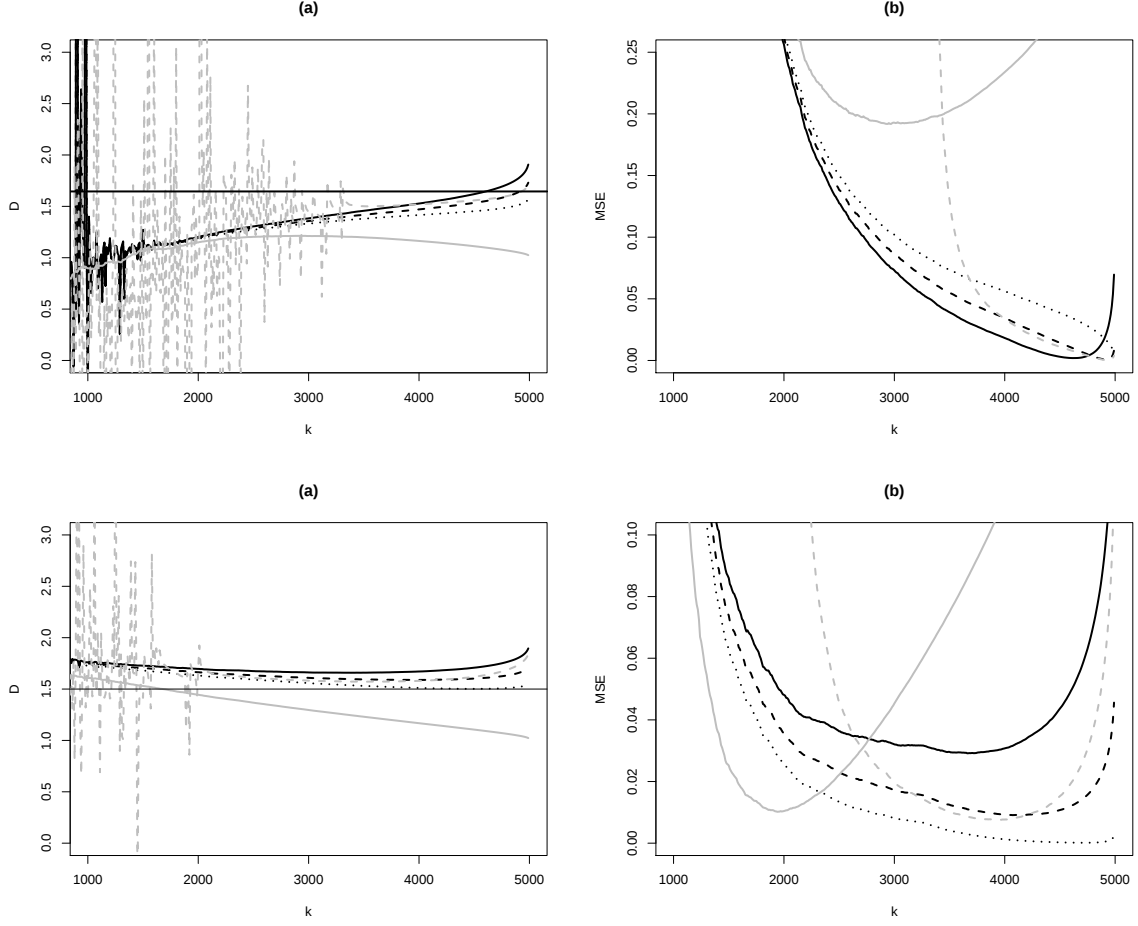


Figure 7: Sample mean (left) and empirical MSE (right) as a function of k when $n = 5000$. $|T_1|$ (top) and $|T_2|$ (bottom) distributions: $\hat{D}_{n,k}(\mathbb{P}^{(1,2)}, \alpha_1, \hat{\rho}_k)$ with $\alpha_1 = 1.5$ (black solid line), 1.75 (black dashed line) and 2.0 (black dotted line), and $\hat{D}_{n,k}^{(CG, \tau)}(\hat{\rho}_k)$ with $\tau = -0.25$ (grey dashed line). The grey solid line represents $\hat{D}_{n,k}^{(GM)}(\hat{\rho}_k)$.

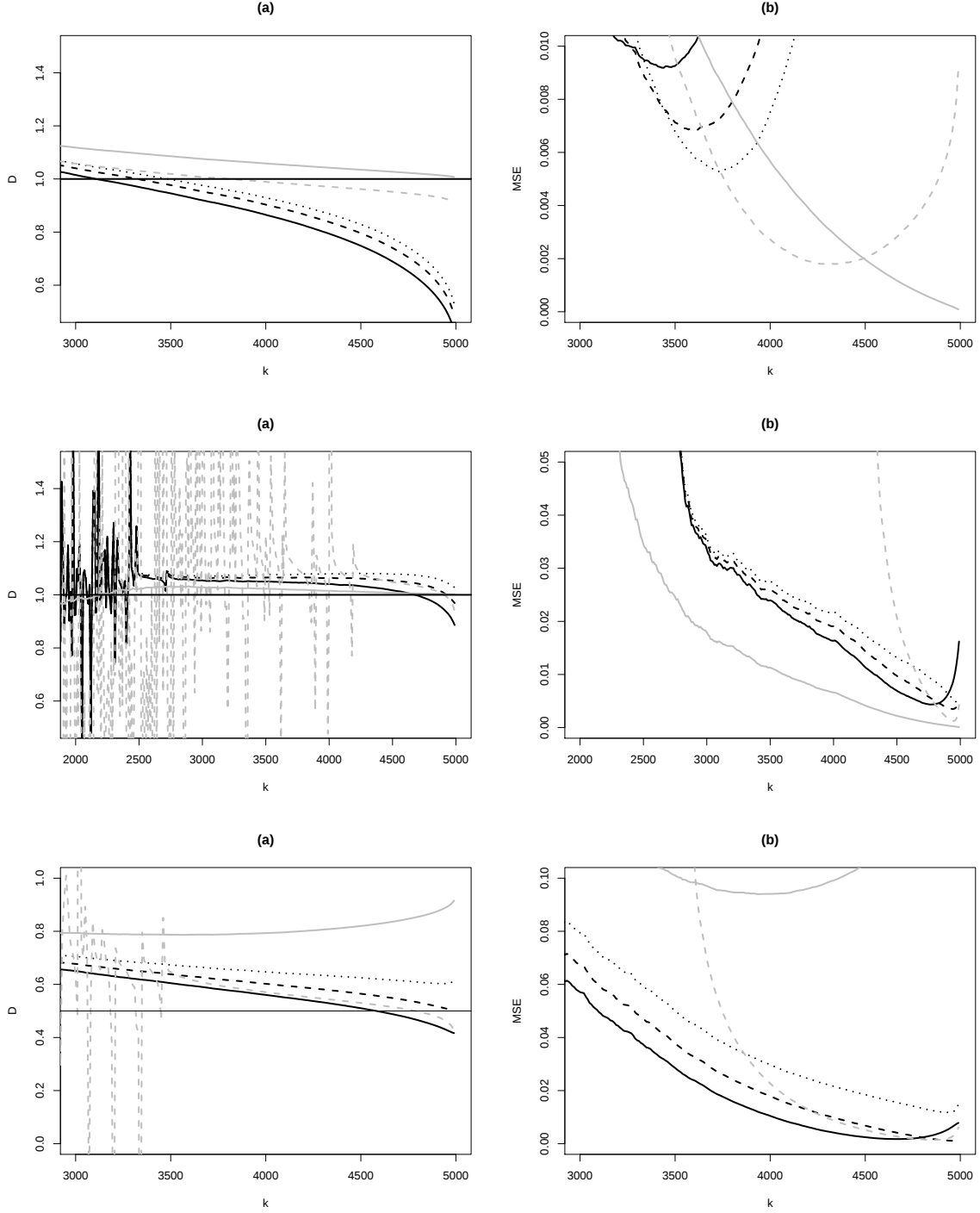


Figure 8: Sample mean (left) and empirical MSE (right) as a function of k when $n = 5000$. Burr(1,1,1) (top) and Burr(1,0.5,2) (middle) distributions: $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \hat{\rho}_k)$ with $\alpha_1 = -3.25$ (black solid line), -3.1 (black dashed line) and -3.0 (black dotted line), and $\hat{D}_{n,k}^{(CG,\tau)}(\hat{\rho}_k)$ with $\tau = -1.2$ (grey dashed line). Fréchet(1) distribution (bottom): $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \hat{\rho}_k)$ with $\alpha_1 = -5.5$ (black solid line), -5.0 (black dashed line) and -4.5 (black dotted line), and $\hat{D}_{n,k}^{(CG,\tau)}(\hat{\rho}_k)$ with $\tau = -3$ (grey dashed line). The grey solid line represents $\hat{D}_{n,k}^{(GM)}(\hat{\rho}_k)$.

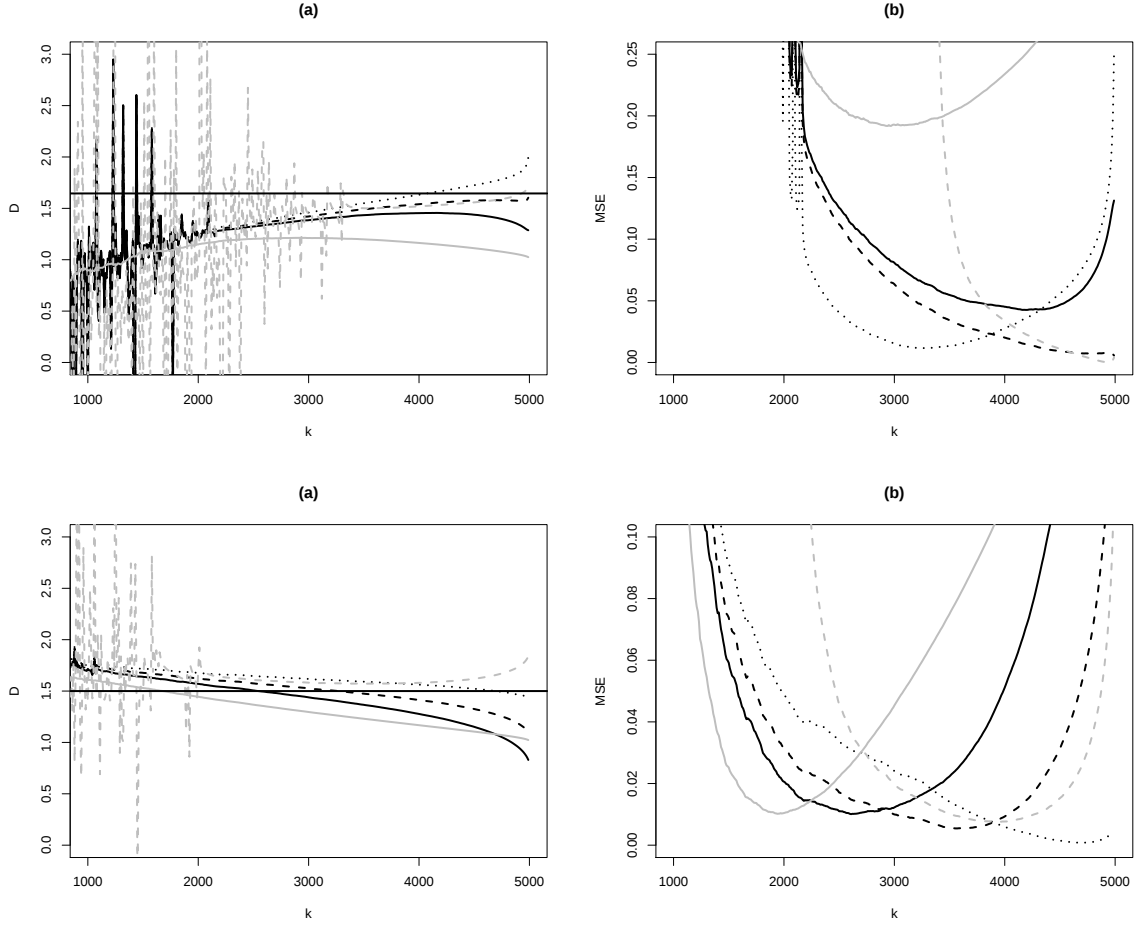


Figure 9: Sample mean (left) and empirical MSE (right) as a function of k when $n = 5000$. $|T_1|$ (top) and $|T_2|$ (bottom) distributions: $\hat{D}_{n,k}(\mathbb{L}^{(1,2)}, \alpha_1, \hat{\rho}_k)$ with $\alpha_1 = -2.0$ (black solid line), -1.75 (black dashed line) and -1.5 (black dotted line), and $\hat{D}_{n,k}^{(CG, \tau)}(\hat{\rho}_k)$ with $\tau = -0.25$ (grey dashed line). The grey solid line represents $\hat{D}_{n,k}^{(GM)}(\hat{\rho}_k)$.

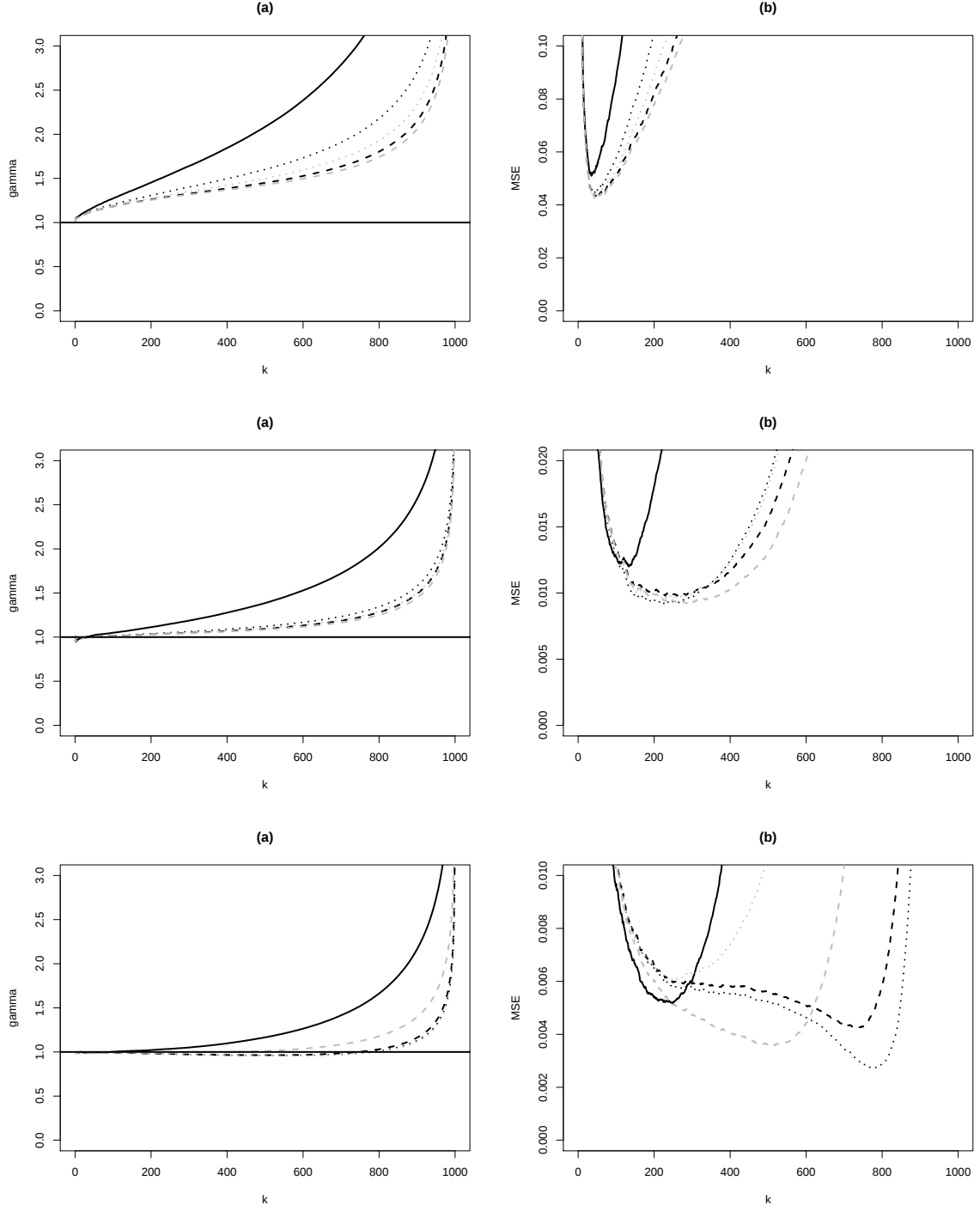


Figure 10: (a) Sample mean and (b) empirical MSE of $H_{k,n}$ (black solid line), and $H_{k,n}^{(BC)}$ implemented with $\hat{D}_{n,\tilde{k}}(\mathbb{P}^{(1,2)}, \alpha_1, \hat{\rho}_{\tilde{k}})$ (black dashed line), $\hat{D}_{n,\tilde{k}}(\mathbb{L}^{(1,2)}, \alpha_1, \hat{\rho}_{\tilde{k}})$ (black dotted line), $\hat{D}_{n,\tilde{k}}^{(GM)}(\hat{\rho}_{\tilde{k}})$ (grey dashed line), and $\hat{D}_{n,\tilde{k}}^{(CG,\tau)}(\hat{\rho}_{\tilde{k}})$ (grey dotted line), as a function of k . Top: Burr(1,2,0.5), Middle: Burr(1,1,1), Bottom: $|T_1|$

Table 1: Examples of Pareto-type distributions satisfying $(\tilde{\mathcal{R}})$

Distribution	$1 - F(x)$	γ	$b(x)$	$\tilde{b}(x)$
Burr(ζ, λ, δ)	$\left(\frac{\zeta}{\zeta+x^\delta}\right)^\lambda$ $x > 0; \zeta, \lambda, \delta > 0$	$\frac{1}{\lambda\delta}$	$\gamma x^{-1/\lambda}$	$x^{-1/\lambda}$
F_{ν_1, ν_2}	$\frac{\Gamma((\nu_1+\nu_2)/2)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} \int_x^\infty u^{\nu_1/2-1} \left(1 + \frac{\nu_1}{\nu_2}u\right)^{-(\nu_1+\nu_2)/2} du$ $x > 0; \nu_1, \nu_2 > 0$	$\frac{2}{\nu_2}$	$\frac{2D_1}{\nu_2} x^{-2/\nu_2}$ $D_1 := \frac{2}{\nu_2} C_1 C_0^{-2/\nu_2}$ $C_0 := \frac{\Gamma((\nu_1+\nu_2)/2)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2+1)} \left(\frac{\nu_1}{\nu_2}\right)^{-\nu_2/2}$ $C_1 := \frac{\nu_2^2(\nu_1+\nu_2)}{2\nu_1(\nu_2+2)}$	$\left(D_1 - 2\frac{D_2}{D_1}\right) x^{-2/\nu_2}$ $D_2 := \left(\frac{2C_2}{\nu_2} - \frac{\nu_2+2}{\nu_2^2} C_1^2\right) C_0^{-4/\nu_2}$ $C_2 := \frac{\nu_2(\nu_1+\nu_2)(\nu_1+\nu_2+2)}{8(\nu_2+4)} \left(\frac{\nu_2}{\nu_1}\right)^2$
Fréchet(α)	$1 - \exp(-x^{-\alpha})$ $x > 0; \alpha > 0$	$\frac{1}{\alpha}$	$\frac{\gamma}{2x}$	$\frac{5}{6x}$
GPD(κ, σ)	$\left(1 + \kappa \frac{x}{\sigma}\right)^{-1/\kappa}$ $x > 0; \kappa, \sigma > 0$	κ	$\gamma x^{-\gamma}$	$x^{-\gamma}$
$ T_\nu $	$\frac{2\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)} \int_x^\infty \left(1 + \frac{u^2}{\nu}\right)^{-(\nu+1)/2} du$ $x > 0; \nu > 0$	$\frac{1}{\nu}$	$\frac{2D_1}{\nu} x^{-2/\nu}$ $D_1 := \frac{1}{\nu} C_1 C_0^{-2/\nu}$ $C_0 := \frac{2\nu^{(\nu-1)/2}\Gamma((\nu+1)/2)}{\sqrt{\nu\pi}\Gamma(\nu/2)}$ $C_1 := \frac{\nu^2(\nu+1)}{2(\nu+2)}$	$\left(D_1 - 2\frac{D_2}{D_1}\right) x^{-2/\nu}$ $D_2 := \left(\frac{C_2}{\nu} - \frac{\nu+3}{2\nu^2} C_1^2\right) C_0^{-4/\nu}$ $C_2 := \frac{\nu^3(\nu+1)(\nu+3)}{8(\nu+4)}$