# Asymptotically unbiased estimator of the extreme value index under random censoring

Martin Bladt, Yuri Goegebeur, Armelle Guillou

# Asymptotically unbiased estimator of the extreme value index under random censoring

Martin Bladt[(1)], Yuri Goegebeur[(2)] and Armelle Guillou[(3)]

[(1)] Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark

[(2)] Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

[(3)] Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France

**Abstract**

We consider bias-corrected estimation of the extreme value index of a Pareto-type distribution in the censoring framework. The initial estimator is based on a Kaplan–Meier integral from which we remove the bias under a second-order framework. This estimator depends on a suitable external estimation of second-order parameters which is also discussed. The weak convergence of the bias-corrected estimator is established. It has the nice property to have the same asymptotic variance as the initial estimator. This nice feature is illustrated on a simulation study where our estimator is compared to alternatives already introduced in the literature. Finally, our methodology is applied on an insurance dataset.

## 1 Introduction

Let $X_1, \ldots, X_n$ be a sample of independent and identically distributed (i.i.d.) random variables with distribution function $F$ of Pareto-type, i.e., such that for some $\gamma_F > 0$

$$1 - F(x) = x^{-\frac{1}{\gamma_F}} \ell_F(x), x > 0, \tag{1}$$

where $\ell_F$ denotes a slowly varying function at infinity, i.e., a positive measurable function such that

$$\frac{\ell_F(tx)}{\ell_F(t)} \longrightarrow 1 \text{ as } t \to \infty \text{ for all } x > 0.$$

The parameter $\gamma_F$ is called the extreme value index and it governs the tail behavior, with larger values indicating heavier tails. This model (1) is completely equivalent to the following model for the tail quantile function $U_F(x) := \inf\{y : F(y) \geqslant 1 - 1/x\}, x > 1$ :

$$U_F(x) = x^{\gamma_F} \ell_U(x), \tag{2}$$

where $\ell_U$ is also a slowly varying function at infinity.

Several estimators of $\gamma_F$ have been proposed in the literature, among them the Hill estimator (see [23]), defined as

$$H_X(k) = \frac{1}{k} \sum_{i=1}^{k} \log X_{n-i+1,n} - \log X_{n-k,n},$$

where $X_{1,n} \leqslant \ldots \leqslant X_{n,n}$ denote the order statistics associated to the $X-$sample and $k = k_n$ is an intermediate sequence, i.e., a sequence such that

$$k_n \to \infty \text{ and } k_n/n \to 0.$$

This estimator is the most famous one since it can be interpreted in different ways, e.g., as an estimator of the slope in a Pareto quantile plot, or also as the uniformly minimum variance unbiased estimator of $\gamma_F$ based on the $(k+1)-$upper order statistics of a sample in case of a strict Pareto model (i.e., a model with $\ell_F(x) = 1$ in (1)), see [1].

Under the first-order framework (2) and if $k$ is an intermediate sequence, we can prove that $H_X(k)$ is consistent in probability for the estimation of $\gamma_F$. However, if we want to establish asymptotic normality of $H_X(k)$, assuming Model (2) is not enough, and we need a second-order condition which measures the rate of convergence of $U_F(tx)/U_F(t)$ towards its limit $x^{\gamma_F}$. There are different types of second-order conditions, some expressed in terms of $F$, $U_F$ or their logarithms. The one used in this paper can be formulated as follows:

**Condition (SOC).** *There exist a positive or negative function* $A(.)$ *with* $A(t) \to 0$ *as* $t \to \infty$, *and a parameter* $\rho \leqslant 0$ *such that the tail quantile function of* $F$ *satisfies for all* $x > 0$

$$\lim_{t\to\infty} \frac{\frac{U_F(tx)}{U_F(t)} - x^{\gamma_F}}{A(t)} = x^{\gamma_F}\, h_\rho(x),$$

*with*

$$h_\rho(x) := \int_1^x u^{\rho-1}\, du = \begin{cases} \frac{x^\rho - 1}{\rho} & \text{if } \rho \neq 0 \\ \log x & \text{otherwise.} \end{cases}$$

Under Condition (SOC) and for an intermediate sequence $k$, we have the following asymptotic representation

$$H_X(k) \stackrel{d}{=} \gamma_F + \gamma_F \frac{N_k}{\sqrt{k}} + \frac{A(\frac{n}{k})}{1-\rho} + o_{\mathbb{P}}\left(A\left(\frac{n}{k}\right)\right), \tag{3}$$

where $N_k$ follows asymptotically a standard normal distribution. Equation (3) shows that $H_X(k)$ is asymptotically normally distributed with variance $\gamma_F^2/k$ and a dominant asymptotic bias term given by $A(n/k)/(1-\rho)$. Since under Condition (SOC), $|A(.)|$ is regularly varying with index $\rho$, this means that as $k$ increases, the bias also increases, while the variance decreases and vice versa. Thus, the choice of $k$ is a question of trade-off between bias and variance. Also, the closer $\rho$ is to 0, the more prominent the bias.

This bias problem in the Hill estimator or other tail parameters' estimators has been extensively studied in the literature, e.g., by [3], [13], [18] or [25]. They proposed second-order reduced-bias extreme value index estimators, but at the cost of an increase in variance, larger than or equal to $[\gamma_F(1-\rho)]^2/(\rho^2\, k)$ when $\rho < 0$. To overcome this drawback, [9], [16] and [19] have introduced estimators which present the advantage to reduce the bias without increasing the asymptotic variance, which is kept, e.g., for the Hill estimator at $\gamma_F^2/k$. They are called minimum-variance reduced-bias (MVRB) estimators and due to these nice properties, they obviously outperform the original estimator (i.e., without bias-correction), in terms of mean squared error. Those estimators are defined when the function $A(.)$ in Condition (SOC) is of the form $A(t) = \gamma_F\, \beta\, t^\rho$ with $(\beta, \rho) \in (\mathbb{R}, \mathbb{R}^-)$ suitably estimated externally. This particular form of the function $A(.)$ is not really restrictive because it is very usual in extreme value theory to assume instead of

Model (2), the following Hall-type model (see [21], [22]):

$$U_F(t) = C\, t^{\gamma_F} \left\{ 1 + \gamma_F \beta\, \frac{t^\rho}{\rho} + o\left(t^\rho\right) \right\}, \quad t \to \infty, \tag{4}$$

where $C > 0, \beta \neq 0$ and $\rho < 0$. In that case, Condition (SOC) is satisfied with $A(t) = \gamma_F \beta\, t^\rho$. Note that most of the usual heavy-tailed models belong to this class, e.g., the Fréchet, Burr or generalized Pareto models. The most simple MVRB-type estimator, introduced by [9], is of the form

$$\widehat{\gamma}_{BC}(k) := H_X(k) \left( 1 - \frac{\widehat{\beta}(k_1)}{1 - \widehat{\rho}(k_1)} \left(\frac{n}{k}\right)^{\widehat{\rho}(k_1)} \right), \tag{5}$$

for suitable external estimators of the second-order parameters $(\beta, \rho)$ based on a new intermediate sequence $k_1$ such that $k = o(k_1)$. The specific form of (5) comes from the specific form of the function $A(.)$ together with an estimator of the dominant asymptotic bias term of the Hill estimator which is subtracted from it.

The idea of this paper is to adapt this bias-corrected estimator (5) to the censoring framework. As far as we know, this topic is completely new in the literature, although it has many applications, e.g., in insurance.

The remainder of the paper is organized as follows. In Section 2, after a presentation of two different approaches concerning the construction of an estimator of $\gamma_F$ in the case of censoring, we explain how to derive a bias-corrected estimator in that context. This requires estimators of the second-order parameters $(\beta, \rho)$ adapted to the censoring framework which are also discussed and for which some asymptotic properties are derived. Then, in Section 3, we examine the finite sample performance of our bias-corrected estimator of $\gamma_F$ through a simulation experiment, while in Section 4, we illustrate the method on a real dataset. All the proofs are postponed to Section 5.

## 2 Estimators of tail parameters in the censoring framework

The general random right censoring framework is the following: additionally to the $X-$sample, consider $Y_1, \ldots, Y_n$ a second sample, also based on i.i.d. random variables from some distribution

$G$, and assume that the two samples, $X$ and $Y$, are independent. We only observe $(Z_i, \delta_i)$, $i = 1, \ldots, n$, where $Z_i := \min(X_i, Y_i)$ and $\delta_i := \mathbb{1}_{\{X_i \leqslant Y_i\}}$. If $H$ denotes the distribution function of $Z$, due to the independence between the two samples, we have

$$1 - H(x) = (1 - F(x))(1 - G(x)).$$

A consistent estimator of $F$ is the Kaplan–Meier estimator (see [24]) defined as

$$\mathbb{F}_n(x) = 1 - \prod_{i=1}^{n} \left[1 - \frac{\delta_{[i:n]}}{n - i + 1}\right]^{\mathbb{1}_{\{Z_{i,n} \leqslant x\}}},$$

where $Z_{1,n} \leqslant \ldots \leqslant Z_{n,n}$ are the order statistics of the $Z-$sample and $\delta_{[i,n]}, i = 1, \ldots, n$, are the concomitants, i.e., $\delta_{[i,n]} = \delta_j$ iff $Z_{i,n} = Z_j$.

In that censoring framework, if we assume that $G$ is also of Pareto-type with index $\gamma_G$, then the Hill estimator based on the observations $Z_1, \ldots, Z_n$, i.e., $H_Z(k)$, estimates the extreme value index of $H$, that is $\gamma_H := \frac{\gamma_F \gamma_G}{\gamma_F + \gamma_G}$, but not $\gamma_F$. To solve this issue, [4] and [12] proposed to divide this Hill estimator by the proportion of non-censored observations in the $k$-largest observations, i.e., they proposed to consider

$$H^{(c)}(k) := \frac{H_Z(k)}{\frac{1}{k}\sum_{i=1}^{k} \delta_{[n-i+1,n]}}$$

as an estimator of $\gamma_F$.

Very recently, [7] proposed an alternative approach based on Kaplan–Meier integrals of the form

$$S_{k,n}(\varphi) := \int \varphi \, d\mathbb{F}_{k,n},$$

where $\varphi$ is a measurable function, $F-$square integrable and almost everywhere continuous on $[1, \infty)$, and $\mathbb{F}_{k,n}$ is the extreme Kaplan–Meier estimator defined as

$$\mathbb{F}_{k,n}(x) = 1 - \prod_{i=1}^{k} \left[1 - \frac{\delta_{[n-i+1:n]}}{i}\right]^{\mathbb{1}_{\{Z_{n-i+1,n}/Z_{n-k,n} \leqslant x\}}}.$$

In this paper, we consider the particular function $\varphi(x) := \varphi_\ell(x) := (\log x)^\ell$, with $\ell$ a positive integer, and we define

$$\mathbb{M}_n^{(\ell)}(k) := \int_1^\infty \varphi_\ell \, d\mathbb{F}_{k,n},$$

as a consistent estimator in probability for

$$M^{(\ell)} := \int_1^\infty \varphi_\ell \, dF^\circ,$$

with $F^\circ(x) := \left(1 - x^{-\frac{1}{\gamma_F}}\right) \mathbb{1}_{\{x \geqslant 1\}}$ as soon as $\gamma_G > \gamma_F$ (see Theorem 3.2 in [7]). Thus, when $\ell = 1$, $\mathbb{M}_n^{(1)}(k)$ can be interpreted as a Hill-type estimator adapted to the censoring framework.

## 2.1 A bias-corrected estimator of $\gamma_F$

Our aim in this paper is to reduce the bias of $\mathbb{M}_n^{(1)}(k)$, without increasing its variance. To reach that goal, we establish the following asymptotic representation, similar to (3)

$$\begin{aligned} \mathbb{M}_n^{(1)}(k) \quad \stackrel{d}{=} \quad & \gamma_F + \gamma_F \sqrt{\frac{\gamma_G}{\gamma_G - \gamma_F}} \frac{N_k^{(c)}}{\sqrt{k}} + \frac{A\left(\left[1 - F\left(U_H\left(n/k\right)\right)\right]^{-1}\right)}{1-\rho} \\ & + o_{\mathbb{P}}\left(A\left(\left[1 - F\left(U_H\left(\frac{n}{k}\right)\right)\right]^{-1}\right)\right), \end{aligned}$$

when $\gamma_G > \gamma_F$ and where $N_k^{(c)}$ follows asymptotically a standard normal distribution (see Lemma 5.1). This result shows that $\mathbb{M}_n^{(1)}(k)$ has a dominant asymptotic bias term of $A\left(\left[1 - F\left(U_H\left(n/k\right)\right)\right]^{-1}\right)/(1-\rho)$ and an asymptotic variance of $\gamma_F^2 \gamma_G / (\left[\gamma_G - \gamma_F\right] k)$. The bias can be estimated, as soon as we have estimators for $\rho$, $1 - F\left(U_H\left(n/k\right)\right)$ and $A(.)$ adapted to the censoring framework. Concerning $1 - F\left(U_H\left(n/k\right)\right)$, this can be done easily by replacing $F$ by the Kaplan–Meier estimator $\mathbb{F}_n$ and $U_H\left(n/k\right)$ by its empirical counterpart $Z_{n-k,n}$. This leads to the estimator $1 - \mathbb{F}_n\left(Z_{n-k,n}\right)$. It thus remains to find estimators for $(\beta, \rho)$ since we assume Model (4), and therefore the specific form for the function $A(.)$. Proposing such estimators is already a new contribution of the paper. If we denote them by $(\widehat{\beta}^{(c)}(k_1), \widehat{\rho}^{(c)}(k_1))$, then, following the construction idea of [9], we introduce

$$\widehat{\gamma}_{BC}^{(c)}(k) := \mathbb{M}_n^{(1)}(k) \left(1 - \frac{\widehat{\beta}^{(c)}(k_1)}{1 - \widehat{\rho}^{(c)}(k_1)} \left[1 - \mathbb{F}_n(Z_{n-k,n})\right]^{-\widehat{\rho}^{(c)}(k_1)}\right), \tag{6}$$

as our bias-corrected Hill-type estimator adapted to the censoring framework.

Our main result, stated in the following theorem, is the weak convergence of our estimator $\widehat{\gamma}_{BC}^{(c)}(k)$, correctly normalized.

**Theorem 2.1** *Consider Model (4) with $F$ continuous, and assume that $G$ is of Pareto-type with index $\gamma_G > \gamma_F$. Let $k$ and $k_1$ be two intermediate sequences such that $k = o(k_1)$ and*

$$\sqrt{k}\, A\left([1 - F(U_H(n/k))]^{-1}\right) \longrightarrow \lambda \in \mathbb{R}.$$

*Then, for $(\widehat{\beta}^{(c)}(k_1), \widehat{\rho}^{(c)}(k_1))$ any estimators of $(\beta, \rho)$ such that*

$$\log\left(\left[1 - F\left(U_H\left(\frac{n}{k}\right)\right)\right]^{-1}\right)\left(\widehat{\rho}^{(c)}(k_1) - \rho\right) = o_{\mathbb{P}}(1) \tag{7}$$

*and*

$$\frac{\widehat{\beta}^{(c)}(k_1)}{\beta} = 1 + o_{\mathbb{P}}(1), \tag{8}$$

*we have*

$$\sqrt{k}\left(\widehat{\gamma}_{BC}^{(c)}(k) - \gamma_F\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\gamma_G\, \gamma_F^2}{\gamma_G - \gamma_F}\right).$$

Remark that Condition (7) involves both the sequences $k$ and $k_1$.

As is clear from this theorem, $\widehat{\gamma}_{BC}^{(c)}(k)$ is an asymptotically unbiased estimator of $\gamma_F$ in the sense that its limiting distribution has a zero mean whatever the value of $\lambda$. Note also that the variance of $\widehat{\gamma}_{BC}^{(c)}(k)$ is the same as the variance of the original $\mathbb{M}_n^{(1)}(k)$ estimator, given in [7] (see also our Lemma 5.1), which means that our estimator $\widehat{\gamma}_{BC}^{(c)}(k)$ can be seen as a MVRB-type estimator, as defined in [9], but adapted to the censoring framework.

## 2.2 Estimator of $\rho$

In this section, we provide an estimator of $\rho$ and since we want to show its weak convergence, we need to strengthen condition (SOC) into a third-order condition, given by:

**Condition (TOC).** *There exist a positive or negative function $B(.)$ with $B(t) \to 0$ as $t \to \infty$, and a parameter $\xi \leqslant 0$ such that the tail quantile function of $F$ satisfies for all $x > 0$*

$$\lim_{t\to\infty} \frac{1}{B(t)} \left\{ \frac{\frac{U_F(tx)}{U_F(t)} - x^{\gamma_F}}{A(t)} - x^{\gamma_F}\, h_\rho(x) \right\} = x^{\gamma_F}\, h_{\rho+\xi}(x).$$

The estimation of the second-order parameter $\rho$ has been considered in the literature in the case where there is no censoring, see, e.g., [8], [14], [15] or [17]. In particular, in [14], the following estimator has been introduced

$$\widehat{\rho}(k_1) := \frac{3\,T_n^{(\tau)}(k_1) - 3}{T_n^{(\tau)}(k_1) - 3} \qquad \text{provided } T_n^{(\tau)}(k_1) \in [1,3),$$

with $\tau > 0$, where

$$T_n^{(\tau)}(k_1) \;:=\; \frac{[M_n^{(1)}(k_1)]^\tau - [\frac{1}{2}\,M_n^{(2)}(k_1)]^{\frac{\tau}{2}}}{[\frac{1}{2}\,M_n^{(2)}(k_1)]^{\frac{\tau}{2}} - [\frac{1}{6}\,M_n^{(3)}(k_1)]^{\frac{\tau}{3}}}$$

with, for $\alpha \in \mathbb{R}^+$,

$$M_n^{(\alpha)}(k_1) \;:=\; \frac{1}{k_1}\sum_{i=1}^{k_1}[\log X_{n-i+1,n} - \log X_{n-k_1,n}]^\alpha.$$

Note that $M_n^{(1)}(k_1)$ is nothing else than the Hill estimator $H_X(k_1)$.

Under Condition (SOC) with $\rho < 0$ and with an intermediate sequence $k_1$ such that $\sqrt{k_1}\,A(n/k_1) \to \infty$, [14] proved the consistency in probability of $\widehat{\rho}(k_1)$ (see their Theorem 2.1). Also, under a third-order condition similar to Condition (TOC) with $\rho < 0$ but expressed in terms of $\log \frac{U_F(tx)}{U_F(t)}$ instead of $\frac{U_F(tx)}{U_F(t)}$, as well as $\sqrt{k_1}\,A(n/k_1) \to \infty$, $\sqrt{k_1}\,A^2(n/k_1) \to \lambda_1 \in \mathbb{R}$ and $\sqrt{k_1}\,A(n/k_1)\,B(n/k_1) \to \lambda_2 \in \mathbb{R}$, their estimator $\widehat{\rho}(k_1)$, correctly normalized, is asymptotically normally distributed (see Theorem 3.1 in [14]).

Our aim in this section is to adapt this estimator to the censoring framework by introducing

$$\widehat{\rho}^{(c)}(k_1) := \frac{3\,\mathbb{T}_n^{(\tau)}(k_1) - 3}{\mathbb{T}_n^{(\tau)}(k_1) - 3} \qquad \text{provided } \mathbb{T}_n^{(\tau)}(k_1) \in [1,3),$$

with

$$\mathbb{T}_n^{(\tau)}(k_1) \;:=\; \frac{[\mathbb{M}_n^{(1)}(k_1)]^\tau - [\frac{1}{2}\,\mathbb{M}_n^{(2)}(k_1)]^{\frac{\tau}{2}}}{[\frac{1}{2}\,\mathbb{M}_n^{(2)}(k_1)]^{\frac{\tau}{2}} - [\frac{1}{6}\,\mathbb{M}_n^{(3)}(k_1)]^{\frac{\tau}{3}}}.$$

The asymptotic normality of our estimator $\widehat{\rho}^{(c)}(k_1)$ is given in the following theorem:

**Theorem 2.2** *Assume that $F$ and $G$ are of Pareto-type with index $\gamma_F$ and $\gamma_G$, respectively, such that $\gamma_G > \gamma_F$. Under the third-order condition (TOC) with $\rho, \xi < 0$, for an intermediate sequence $k_1$ satisfying*

$$\begin{aligned}
&\sqrt{k_1}\, A\left(\left[1 - F\left(U_H\left(n/k_1\right)\right)\right]^{-1}\right) \longrightarrow \infty, \\
&\sqrt{k_1} A^2\left(\left[1 - F\left(U_H\left(n/k_1\right)\right)\right]^{-1}\right) \longrightarrow \lambda_1 \in \mathbb{R}, \\
&\sqrt{k_1} A\left(\left[1 - F\left(U_H\left(n/k_1\right)\right)\right]^{-1}\right)\, B\left(\left[1 - F\left(U_H\left(n/k_1\right)\right)\right]^{-1}\right) \longrightarrow \lambda_2 \in \mathbb{R},
\end{aligned}$$

*we have*

$$\sqrt{k_1}\, A\left(\left[1 - F\left(U_H\left(\frac{n}{k_1}\right)\right)\right]^{-1}\right)\left\{\widehat{\rho}^{(c)}(k_1) - \rho\right\} \xrightarrow{d} \mathcal{N}\left(m, \sigma^2\right),$$

*where*

$$\begin{aligned}
m &:= \left\{-\frac{2(1-\rho)^3}{(1-2\rho)^3} + \frac{\rho^2(7 - 20\rho + 16\rho^2 - 4\rho^3)}{2(1-2\rho)^2(1-\rho)^2}\frac{1}{\gamma_F} + \frac{\rho(3-2\rho)(3-\rho)}{12(1-\rho)^2}\frac{\tau}{\gamma_F}\right\}\lambda_1 \\
&\quad + \frac{\xi(1-\rho)^3(\rho+\xi)}{\rho(1-\rho-\xi)^3}\,\lambda_2 \\
\sigma^2 &:= \frac{\gamma_G\,\gamma_F^2\,(1-\rho)^6}{\rho^2(\gamma_G - \gamma_F)^5}\left\{(2\rho^2 - 2\rho + 1)\gamma_G^4 - 2\rho(2-\rho)\gamma_G^3\gamma_F\right. \\
&\qquad \left. + 2(\rho^2 - 2\rho + 2)\gamma_G^2\gamma_F^2 - 2\rho\gamma_G\gamma_F^3 + \gamma_F^4\right\}.
\end{aligned}$$

Note that in the particular case where $\gamma_G \to \infty$ we recover the variance in the uncensored case, see (3.18) in [14]. Remark also that this estimator satisfies Condition (7) of our Theorem 2.1 as soon as

$$\frac{\log\left(\left[1 - F\left(U_H\left(\frac{n}{k}\right)\right)\right]^{-1}\right)}{\sqrt{k_1}\, A\left(\left[1 - F\left(U_H\left(\frac{n}{k_1}\right)\right)\right]^{-1}\right)} = o(1). \tag{9}$$

### 2.3 Estimator of $\beta$

Now, we discuss the estimation of $\beta$ adapted to the censoring framework. Using a similar approach as for the estimation of $\rho$, we propose the following estimator of $\beta$

$$\widehat{\beta}^{(c)}(k_1) := \frac{\kappa^{\frac{1}{\kappa-1}}\,\tau^{-1}}{\frac{1}{\theta_1\,\widehat{\rho}^{(c)}(k_1)}\left\{\left(\frac{1}{1-\widehat{\rho}^{(c)}(k_1)}\right)^{\ell\,\theta_1}-1\right\}-\frac{1}{\theta_2\,\widehat{\rho}^{(c)}(k_1)}\left\{\left(\frac{1}{1-\widehat{\rho}^{(c)}(k_1)}\right)^{\ell\,\theta_2}-1\right\}} \times\left(\frac{\left[\left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{\Gamma(\ell\,\theta_1+1)}\right)^{\frac{\tau}{\theta_1}}-\left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{\Gamma(\ell\,\theta_2+1)}\right)^{\frac{\tau}{\theta_2}}\right]^{\kappa}}{\left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{\Gamma(\ell\,\theta_1+1)}\right)^{\frac{\tau}{\theta_1}\,\kappa}-\left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{\Gamma(\ell\,\theta_2+1)}\right)^{\frac{\tau}{\theta_2}\,\kappa}}\right)^{\frac{1}{\kappa-1}} [1-\mathbb{F}_n(Z_{n-k_1,n})]^{\widehat{\rho}^{(c)}(k_1)}.$$

Since we only need consistency in probability in Condition (8), we work under Model (4) instead of Condition (TOC).

**Theorem 2.3** *Consider Model (4) with $F$ continuous, and assume that $G$ is of Pareto-type with index $\gamma_G > \gamma_F$. For an intermediate sequence $k_1$ satisfying*

$$\frac{\log\left(\left[1-F\left(U_H\left(\frac{n}{k_1}\right)\right)\right]^{-1}\right)}{\sqrt{k_1}\,A\left(\left[1-F\left(U_H\left(\frac{n}{k_1}\right)\right)\right]^{-1}\right)}=o(1), \tag{10}$$

*we have*

$$\frac{\widehat{\beta}^{(c)}(k_1)}{\beta}\xrightarrow{\mathbb{P}}1.$$

Note that (10) is similar to (9), and since $k=o(k_1)$ in Theorem 2.1, if (9) is satisfied, this is also the case for (10). In other words if $\widehat{\gamma}_{BC}^{(c)}(k)$ in (6) is based on the estimators of $(\beta,\rho)$ defined in Sections 2.2 and 2.3, then Conditions (7) and (8) can be simply replaced by Condition (9).

## 3 Simulation study

We consider a simulation study where $X$ and $Y$ follow a Burr distribution whose distribution function is of the form

$$D(x)=1-\left(1+x^{\vartheta}\right)^{-\lambda},x>0, \tag{11}$$

where $\lambda > 0, \vartheta > 0$ and $D$ is either $F$ or $G$. In that case, $F$ satisfies the Hall-type model (4) with an extreme value index $\gamma_F = 1/(\lambda\,\vartheta)$ and a second order parameter $\rho = -1/\lambda$. We consider three Burr distributions for $X$ corresponding to $(\lambda, \vartheta)$ in (11) equal to $(1/2, 4)$, $(1, 2)$ and $(2, 1)$. This gives the same tail index $\gamma_F = 0.5$ and a second order parameter $\rho = -2, -1, -0.5$, respectively. For the censoring variable $Y$, also three Burr distributions are considered with parameters $(1/9.5, 1)$, $(1/4.5, 1)$ and $(1/2, 1)$, leading to tail indices $\gamma_G = 9.5, 4.5, 2$, respectively, and thus an asymptotic proportion of censoring, given by $\gamma_F/(\gamma_F + \gamma_G)$, of $5\%, 10\%$ and $20\%$, respectively.

We consider $N = 1\,000$ simulation replicates of size $n = 500$. For each of the nine combinations, we construct three preliminary plots of the mean of $\widehat{\rho}^{(c)}(k_1)$ as a function of $k_1$, one for each $\tau = 0.5, 1, 2$, and we choose first the $\tau$ that provides the most stability and then select the largest $k_1$ value within the stable region. This provides nine chosen tuples of $(\tau, k_1)$, which are then considered fixed in the remainder of the study. We then compute the bias-reduced Hill estimator adapted to the censoring framework, given by $\widehat{\gamma}^{(c)}_{BC}(k)$ and defined in (6), and where the estimator of $\beta$ is evaluated with $(\kappa, \ell, \theta_1, \theta_2) = (2, 1, 1, 2)$. We compare our estimator with two alternative ones: the estimator $\mathbb{M}^{(1)}_n(k)$ proposed in [7] which is not asymptotically unbiased, and a bias-corrected estimator proposed in [2] in case a parameter, denoted $\beta^*$ (given under model (4) as the smallest value of the $\vartheta$-parameters between the distributions of $F$ and $G$), is assumed known. Note that bias-correction is achieved by this estimator only when $\beta^*$ is known (see Theorem 1 in [2]), but since, as far as we know, this estimator is the only one already existing in the literature which is asymptotically unbiased in the context of censoring, we used it here. In our plots, since $\beta^*$ is unknown, we follow the recommendation given in [2] by replacing it by $\rho^* k/(\sum_{i=1}^{k} \log Z_{n-i+1,n} - \log Z_{n-k,n})$, with $\rho^* = -1/2, -1, -2$. Then, the best performing of the three choices is shown in our figures.

The comparison between the three estimators is shown in Figures 1, 2 and 3 for the three different Burr distributions for $X$, where the mean and mean squared error (MSE) of the three estimators of $\gamma_F$ are plotted as a function of $k$, for $k$ between $0.05\,n$ and $k_1$. This guarantees the constraint $k \leqslant k_1$. The horizontal line on the left panel of each figure corresponds to the true value of $\gamma_F$ and the three rows to the three Burr distributions of $Y$, leading to the asymptotic

proportion of censoring $5\%, 10\%$ and $20\%$, from the top to the bottom.

As is well-known in extreme value theory, if $\rho$ is close to zero, the bias of the extreme value index estimator can be large. On the contrary, when $|\rho|$ increases, it tends to become less severe. Since in Figure 1, $\rho = -2$ which is far away from 0, we expect almost no bias for $\mathbb{M}_n^{(1)}(k)$, and thus bias-correction will not improve the estimator considerably. This is exactly what we observe in Figure 1, where roughly the three estimators work similarly in terms of minimum MSE, with a better behavior for the estimator defined in [2] in terms of bias. In Figure 2, $\rho = -1$ which is still not too close to 0, but nevertheless a bias already appears for $\mathbb{M}_n^{(1)}(k)$, which increases quickly with $k$. In that case bias reduction is useful and our estimator $\widehat{\gamma}_{BC}^{(c)}(k)$ performs very well in terms of bias and MSE, with a long stability of the sample paths as a function of $k$. It is clearly the best estimator among the three. In Figure 3, $\rho = -0.5$, which is close to 0 and in that case our estimator $\widehat{\gamma}_{BC}^{(c)}(k)$ outperforms drastically the two alternatives, in terms of bias for the $\mathbb{M}_n^{(1)}(k)$ estimator, but more importantly in terms of MSE against the two alternatives, with again a long stability as a function of $k$. This is a nice feature of our estimator, since in that case the choice of $k$ is not so crucial.

In our simulation, we observed that the method of [2] is rather sensitive to the choice of the $\rho^*$ hyperparameter, which can also be seen in Figure 2 of [2], while our method is significantly more insensitive to the choice of $(\tau, k_1)$.

# 4 Real data application

We consider a dataset of 109 992 claims settlement observations corresponding to a damage guarantee of a French issuer from 1992 to 2007, publicly available through the `R`-package `CASdatasets`, listed as the dataset `freclaimset3dam9207`. Some observations are right censored since claims are not fully settled at the date of data collection. This may happen, for instance, if legal proceedings are ongoing. The dataset is illustrated in Figure 4 (top left) where log-claim sizes above 10k€ are depicted in chronological order, closed claims being the black circles while open claims are the red triangles. As expected, the more recently arrived claims show more censoring than older claims. Note that this dataset has already been analyzed in the extreme value literature (see [5] and [6]), but with emphasis on covariates and goodness of fit,

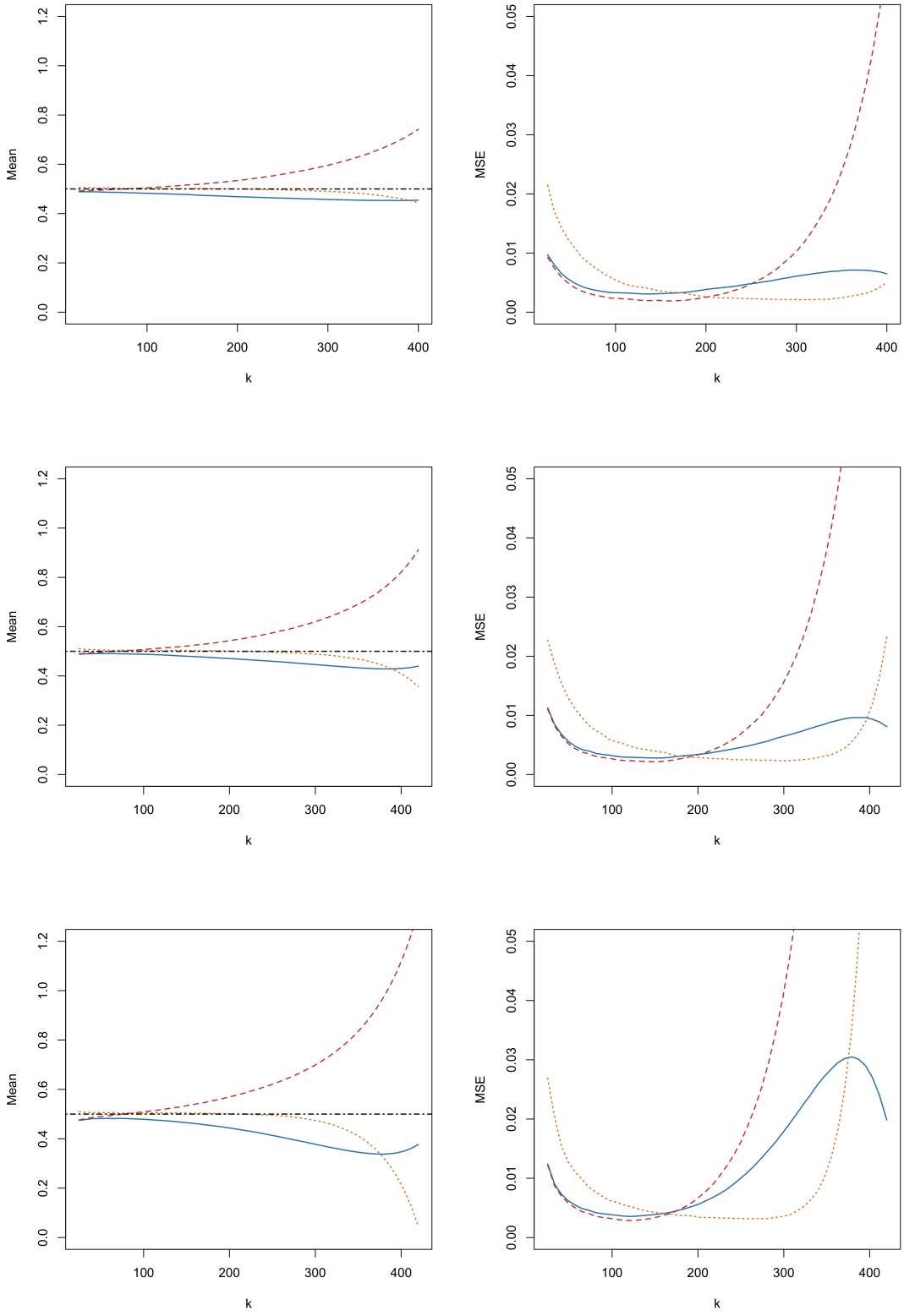




Figure 1: Mean (left panels) and MSE (right panels) of $\widehat{\gamma}_{BC}^{(c)}(k)$ (solid blue line), $\mathbb{M}_n^{(1)}(k)$ (dashed red line) and the estimator from [2] (dotted orange line) based on $N = 1\,000$ simulations of size $n = 500$, as a function of $k$. The parameters of the Burr distribution of $X$ are $(1/2, 4)$ while those for $Y$ correspond to 5%, 10% and 20% of censoring, from the top to the bottom panels.

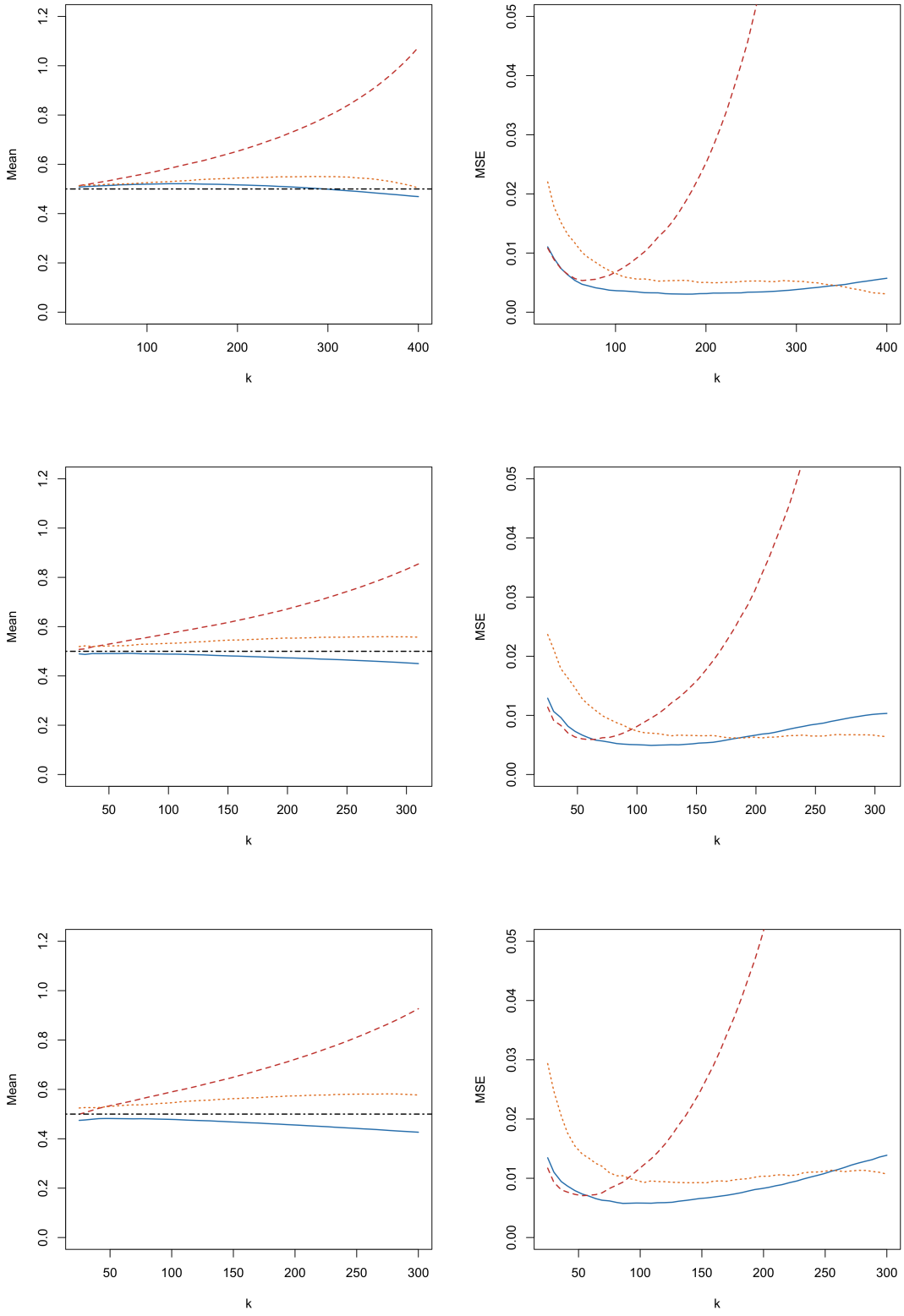




Figure 2: Mean (left panels) and MSE (right panels) of $\widehat{\gamma}_{BC}^{(c)}(k)$ (solid blue line), $\mathbb{M}_n^{(1)}(k)$ (dashed red line) and the estimator from [2] (dotted orange line) based on $N = 1\,000$ simulations of size $n = 500$, as a function of $k$. The parameters of the Burr distribution of $X$ are $(1, 2)$ while those for $Y$ correspond to 5%, 10% and 20% of censoring, from the top to the bottom panels.

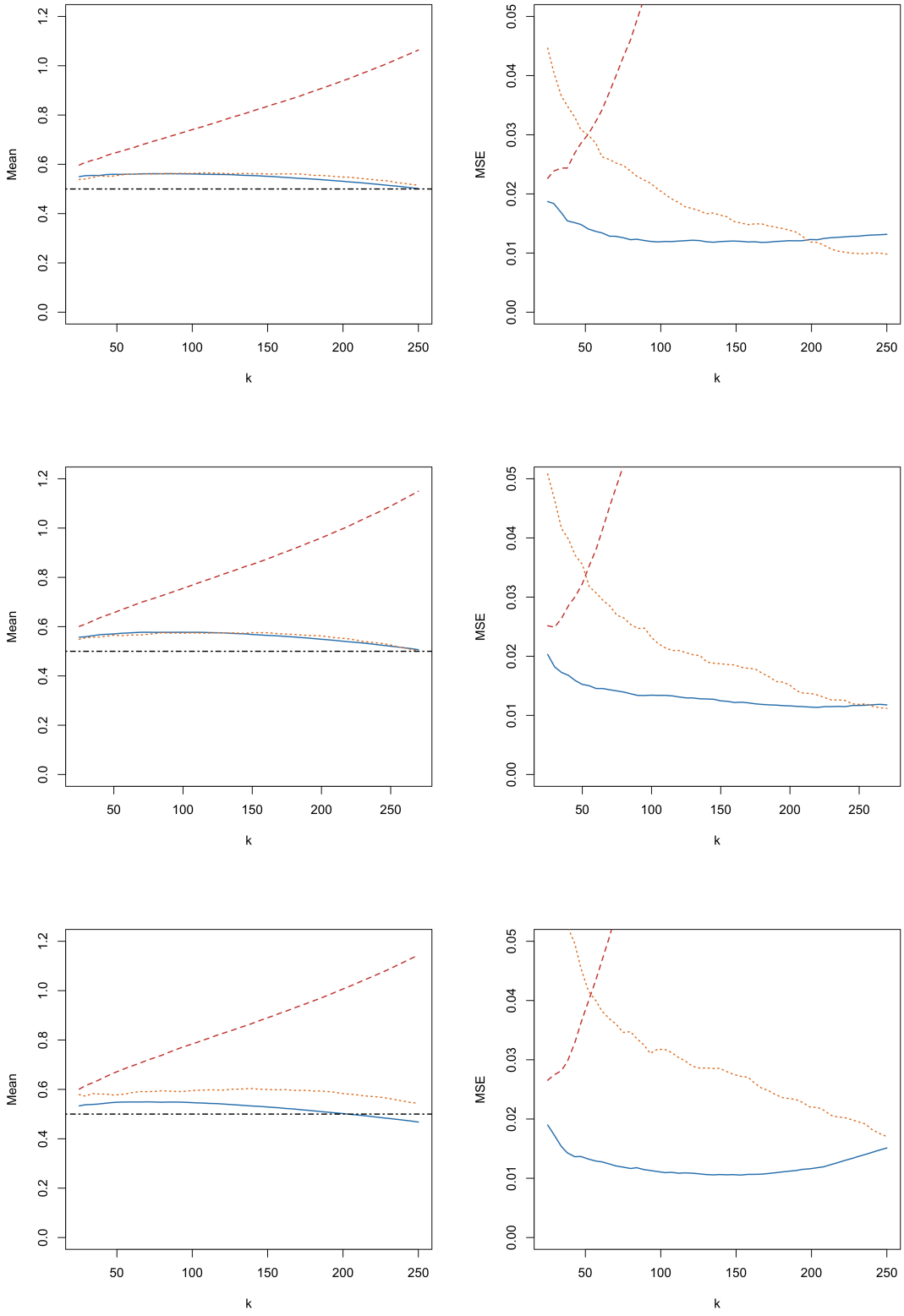




Figure 3: Mean (left panels) and MSE (right panels) of $\widehat{\gamma}_{BC}^{(c)}(k)$ (solid blue line), $\mathbb{M}_n^{(1)}(k)$ (dashed red line) and the estimator from [2] (dotted orange line) based on $N = 1\,000$ simulations of size $n = 500$, as a function of $k$. The parameters of the Burr distribution of $X$ are $(2, 1)$ while those for $Y$ correspond to 5%, 10% and 20% of censoring, from the top to the bottom panels.

instead of bias correction.

To begin with, we check the validity of the assumption that the underlying distribution of the claim sizes is of Pareto-type by plotting on Figure 4 (top right) the Kaplan–Meier Pareto Quantile plot introduced in [4] as

$$\left(-\log\left(1-\mathbb{F}_n(Z_{n-j+1,n})\right), \log Z_{n-j+1,n}\right), j = 1, ..., n-1.$$

The linear pattern in the largest observations confirms an underlying Pareto-type distribution. Then in the bottom of Figure 4, we plot the empirical proportion of non-censored observations in the $k$-largest observations, i.e., $\frac{1}{k}\sum_{i=1}^{k}\delta_{[n-i+1,n]}$ as a function of $k$. This proportion is stable, around 91%, and thus the condition $\gamma_G > \gamma_F$ is clearly satisfied. Next, we estimate $\rho$ similarly as in the simulation study. Namely, we plot $\widehat{\rho}^{(c)}(k_1)$ as a function of $k_1$, for $\tau = 0.5, 1$ and 2, and we select the value of $\tau$ for which we obtain the longest stable region. Then, the selected $k_1-$value is the tipping point where the stability starts to degrade. This strategy is illustrated on Figure 5 (left) and leads to $(\tau, k_1) = (0.5, 19\ 000)$. With these parameter values, we plot on the right panel of Figure 5 our bias-corrected estimator $\widehat{\gamma}_{BC}^{(c)}(k)$ (solid blue line) as a function of $k$, as well as the two alternative estimators $\mathbb{M}_n^{(1)}(k)$ (dashed red line) and the estimator introduced in [2] computed with their recommended algorithm to select $\beta^*$ (dotted orange line). The estimator $\mathbb{M}_n^{(1)}(k)$ has a clearly increasing behavior as a function of $k$ without stability, which indicates a probable presence of bias and a value of $\rho$ also probably close to 0. Regarding the estimator introduced in [2], it is stable only for the second half of the range of values of $k$, after a long increase on the first half. On the contrary, our estimator exhibits long stability for almost all $k$-values at a level which is close to the one where the three estimators were similar, i.e., for very small values of $k$.

# 5 Proofs

## 5.1 Proof of Theorem 2.2.

In the sequel $\zeta$ is a random variable from a strict Pareto(1) distribution, independent of the $Z-$sample. Let $V^\circ$ be a random variable from a strict Pareto($1/\gamma_H$) distribution and $\delta^\circ$ a

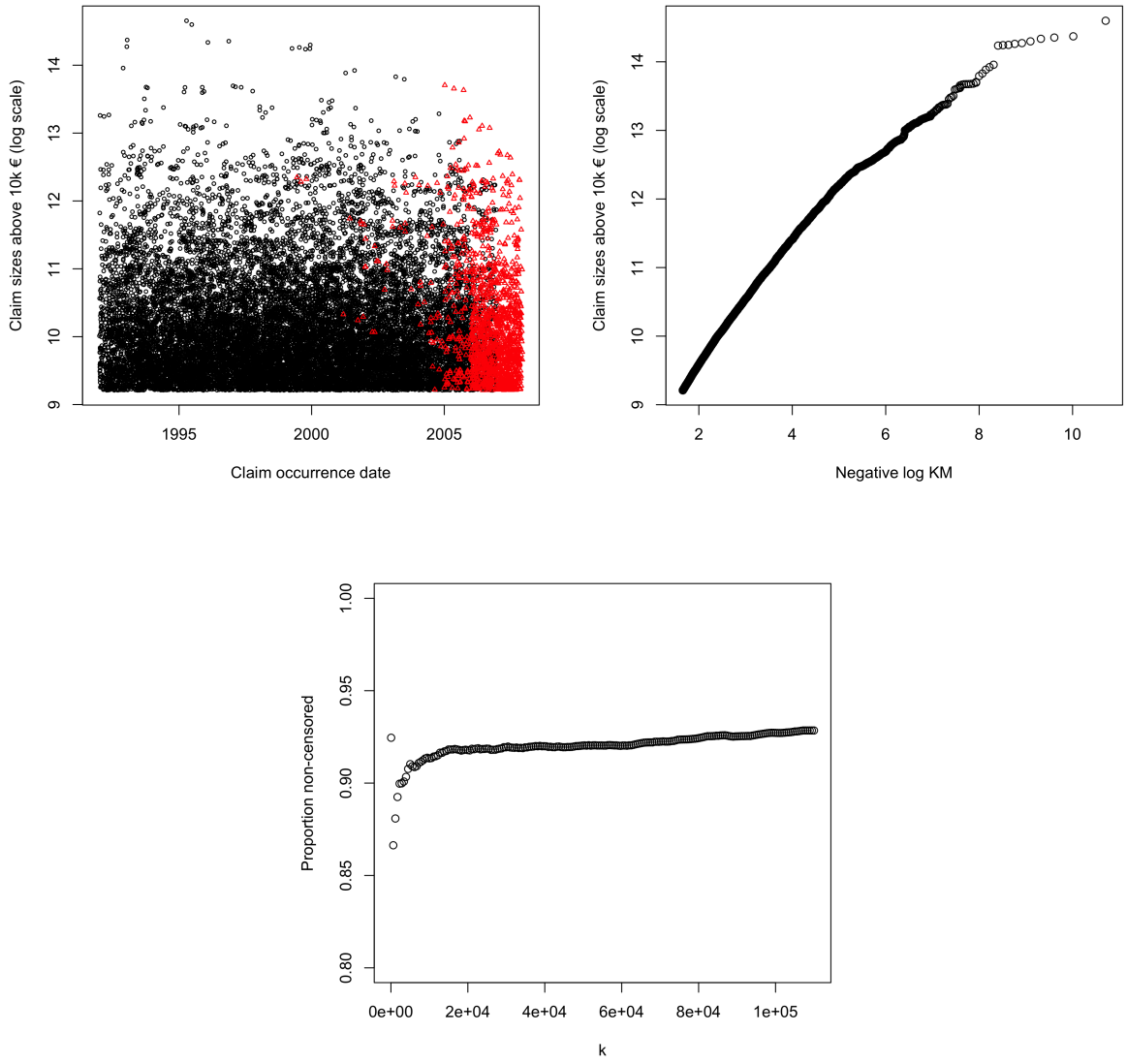


Figure 4: French damage insurance. Top left panel: open claim sizes (red triangles) and closed claim sizes (black circles), top right panel: Kaplan–Meier Pareto quantile plot, bottom panel: empirical proportion of non-censored observations in the $k$-largest observations.

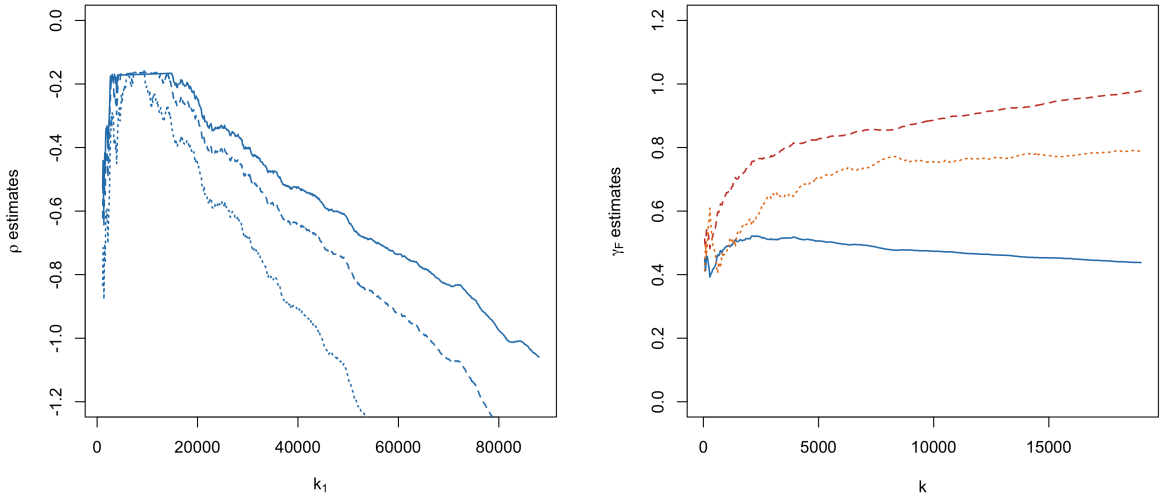


Figure 5: French damage insurance. Left panel: $\widehat{\rho}^{(c)}(k_1)$ as a function of $k_1$, for $\tau = 0.5, 1, 2$ (solid, dashed, dotted lines, respectively). Right panel: $\widehat{\gamma}^{(c)}_{BC}(k)$ (solid blue line), $\mathbb{M}^{(1)}_n(k)$ (dashed red line) and the estimator from [2] (dotted orange line) as a function of $k$.

random variable from a Bernoulli$(\gamma_G/(\gamma_F+\gamma_G))$ distribution, with $V^\circ$ and $\delta^\circ$ independent. We use the same notations as in [7], in particular

$$\begin{aligned}
F^t(x) &:= \frac{F(xt)-F(t)}{1-F(t)}, \quad \text{for } x \geqslant 1,\\
\gamma_0^\circ(x) &:= x^{\frac{1}{\gamma_G}},\\
\gamma_1^\circ(x) &:= \frac{1}{\gamma_F}\, x^{\frac{1}{\gamma_H}} \int_x^\infty \varphi_\ell(z) z^{-\frac{1}{\gamma_F}-1} dz,\\
\gamma_2^\circ(x) &:= \frac{1}{\gamma_F\gamma_G} \int_1^\infty \int_1^\infty \mathbb{1}_{\{v<\min(x,z)\}} \varphi_\ell(z)\, z^{-\frac{1}{\gamma_F}-1}\, v^{\frac{1}{\gamma_H}-1}\, dv\, dz.
\end{aligned}$$

We consider the decomposition

$$\begin{aligned}
\int_1^\infty \varphi_\ell \, d\mathbb{F}_{k_1,n} &= \int_1^\infty \varphi_\ell \, dF^\circ + \int_1^\infty \varphi_\ell \, d\left(\mathbb{F}_{k_1,n} - F^{Z_{n-k_1,n}}\right)\\
&\quad + \int_1^\infty \varphi_\ell \, d\left(F^{Z_{n-k_1,n}} - F^\circ\right), \qquad (12)
\end{aligned}$$

and we define the random variable

$$W^\circ(\varphi_\ell) = \varphi_\ell(V^\circ)\gamma_0^\circ(V^\circ)\delta^\circ + \gamma_1^\circ(V^\circ)(1-\delta^\circ) - \gamma_2^\circ(V^\circ). \qquad (13)$$

According to Theorem 3.4 in [7], when $\gamma_G > \gamma_F$, we have

$$\sigma^{(\ell)}\, G_{k_1}^{(\ell)} := \sqrt{k_1} \int_1^\infty \varphi_\ell \, d\left(\mathbb{F}_{k_1,n} - F^{Z_{n-k_1,n}}\right) \xrightarrow{d} \mathcal{N}(0, [\sigma^{(\ell)}]^2), \tag{14}$$

where

$$[\sigma^{(\ell)}]^2 \;:=\; Var(W^\circ(\varphi_\ell)).$$

Now, using the notations

$$\begin{aligned}
\mu(s) &:= \mathbb{E}[\varphi_\ell(u(s,\zeta))],\\
\mu(\infty) &:= \mathbb{E}[\varphi_\ell(u(\infty,\zeta))],\\
u(s,\zeta) &:= \frac{U_F(s\zeta)}{U_F(s)},\\
u(\infty,\zeta) &:= \zeta^{\gamma_F},
\end{aligned}$$

we have the representation

$$\int_1^\infty \varphi_\ell \, d\left(F^{Z_{n-k_1,n}} - F^\circ\right) = \mu\left([1 - F(Z_{n-k_1,n}-)]^{-1}\right) - \mu(\infty).$$

For a fixed $s$, by a Taylor expansion, we have

$$\begin{aligned}
\mu(s) - \mu(\infty) &= \mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\,\{u(s,\zeta) - u(\infty,\zeta)\}\right]\\
&\quad + \frac{1}{2}\mathbb{E}\left[\varphi_\ell''(u^*(s))\,\{u(s,\zeta) - u(\infty,\zeta)\}^2\right],
\end{aligned} \tag{15}$$

where $u^*(s)$ is an intermediate random value between $u(s,\zeta)$ and $u(\infty,\zeta)$.

To handle $u(s,\zeta) - u(\infty,\zeta)$, we need to use Condition (TOC), which is equivalent to

$$\lim_{t\to\infty} \frac{1}{B(t)} \left\{ \frac{\ell_U(tx) - \ell_U(t)}{A(t)\,\ell_U(t)} - \; h_\rho(x) \right\} = h_{\rho+\xi}(x).$$

Now, according to Proposition 1.3 in [10], there exist functions $a_0$ and $B_0$ such that

$$\lim_{t\to\infty} \frac{1}{B_0(t)} \left\{ \frac{\ell_U(tx) - \ell_U(t)}{a_0(t)} - h_\rho(x) \right\} = h_{\rho+\xi}(x) \tag{16}$$

and for any $\varepsilon, \delta > 0$, there exists a $t_0 = t_0(\varepsilon,\delta) > 0$ such that for all $t, tx \geqslant t_0$

$$\left| \frac{1}{B_0(t)} \left\{ \frac{\ell_U(tx) - \ell_U(t)}{a_0(t)} - \; h_\rho(x) \right\} - h_{\rho+\xi}(x) \right| \leqslant \varepsilon\, x^{\rho+\xi}\, \max\{x^\delta, x^{-\delta}\}. \tag{17}$$

Now (16) is also equivalent to

$$\lim_{t\to\infty}\frac{1}{B_0(t)}\left\{\frac{\frac{U_F(tx)}{U_F(t)}-x^{\gamma_F}}{\frac{a_0(t)}{\ell_U(t)}}-x^{\gamma_F}\,h_\rho(x)\right\}=x^{\gamma_F}\,h_{\rho+\xi}(x),$$

which, in view of Condition (TOC) implies that $A\sim\frac{a_0}{\ell_U}=:A_0$ and $B\sim B_0$.

Now, we consider the decomposition

$$\begin{aligned}u(s,\zeta)-u(\infty,\zeta) &= A_0(s)\zeta^{\gamma_F}\,h_\rho(\zeta)+A_0(s)B_0(s)\,\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\\ &\quad +A_0(s)B_0(s)\left[\frac{1}{B_0(s)}\left\{\frac{\frac{U_F(s\zeta)}{U_F(s)}-\zeta^{\gamma_F}}{A_0(s)}-\zeta^{\gamma_F}\,h_\rho(\zeta)\right\}-\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right].\end{aligned}$$

Combined with (15), this implies that

$$\begin{aligned}&\mu(s)-\mu(\infty)\\ &= A_0(s)\,\mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\,\zeta^{\gamma_F}h_\rho(\zeta)\right]+A_0(s)\,B_0(s)\,\mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\,\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right]\\ &+\frac{1}{2}A_0^2(s)\,\mathbb{E}\left[\varphi_\ell''(u^*(s))\,\zeta^{2\gamma_F}h_\rho^2(\zeta)\right]\\ &+A_0(s)B_0(s)\mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\right.\\ &\qquad\left.\times\left(\frac{1}{B_0(s)}\left\{\frac{\frac{U_F(s\zeta)}{U_F(s)}-\zeta^{\gamma_F}}{A_0(s)}-\zeta^{\gamma_F}\,h_\rho(\zeta)\right\}-\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right)\right]\\ &+A_0^2(s)\,B_0(s)\,\mathbb{E}\left[\varphi_\ell''(u^*(s))\,\zeta^{2\gamma_F}\,h_\rho(\zeta)\,h_{\rho+\xi}(\zeta)\right]\\ &+A_0^2(s)\,B_0(s)\,\mathbb{E}\left[\varphi_\ell''(u^*(s))\,\zeta^{\gamma_F}\,h_\rho(\zeta)\right.\\ &\qquad\left.\times\left(\frac{1}{B_0(s)}\left\{\frac{\frac{U_F(s\zeta)}{U_F(s)}-\zeta^{\gamma_F}}{A_0(s)}-\zeta^{\gamma_F}\,h_\rho(\zeta)\right\}-\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right)\right]\\ &+\frac{1}{2}\,A_0^2(s)\,B_0^2(s)\,\mathbb{E}\left[\varphi_\ell''(u^*(s))\,\zeta^{2\gamma_F}\,h_{\rho+\xi}^2(\zeta)\right]\\ &+\frac{1}{2}\,A_0^2(s)\,B_0^2(s)\,\mathbb{E}\left[\varphi_\ell''(u^*(s))\right.\\ &\qquad\left.\times\left(\frac{1}{B_0(s)}\left\{\frac{\frac{U_F(s\zeta)}{U_F(s)}-\zeta^{\gamma_F}}{A_0(s)}-\zeta^{\gamma_F}\,h_\rho(\zeta)\right\}-\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right)^2\right]\\ &+A_0^2(s)\,B_0^2(s)\,\mathbb{E}\left[\varphi_\ell''(u^*(s))\,\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right.\\ &\qquad\left.\times\left(\frac{1}{B_0(s)}\left\{\frac{\frac{U_F(s\zeta)}{U_F(s)}-\zeta^{\gamma_F}}{A_0(s)}-\zeta^{\gamma_F}\,h_\rho(\zeta)\right\}-\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right)\right]\end{aligned}$$

$$
\begin{aligned}
=: \quad & A_0(s)\,\mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\,\zeta^{\gamma_F} h_\rho(\zeta)\right] + A_0(s)\,B_0(s)\,\mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\,\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right] \\
& + \frac{1}{2} A_0^2(s)\,\mathbb{E}\left[\varphi_\ell''(u^*(s))\,\zeta^{2\gamma_F} h_\rho^2(\zeta)\right] + \sum_{i=1}^{6} T_i(s).
\end{aligned}
$$

We can show that, as $s \to \infty$,

$$
\mathbb{E}\left[\varphi_\ell''(u^*(s))\,\zeta^{2\gamma_F} h_\rho^2(\zeta)\right] \longrightarrow \mathbb{E}\left[\varphi_\ell''(\zeta^{\gamma_F})\,\zeta^{2\gamma_F} h_\rho^2(\zeta)\right], \tag{18}
$$

using the continuity of $\varphi_\ell''$ combined with Potter's inequalities (see, e.g., [20], Proposition B.1.9.5) and the dominated convergence theorem. Now, concerning $T_1(s)$, we can use (17), according which, for all $\varepsilon, \delta > 0$, there exists $s_0 = s_0(\varepsilon, \delta)$ such that $\forall s \geqslant s_0$

$$
\begin{aligned}
& \left| \mathbb{E}\left[ \varphi_\ell'(\zeta^{\gamma_F}) \left( \frac{1}{B_0(s)} \left\{ \frac{\frac{U_F(s\zeta)}{U_F(s)} - \zeta^{\gamma_F}}{A_0(s)} - \zeta^{\gamma_F}\, h_\rho(\zeta) \right\} - \zeta^{\gamma_F}\, h_{\rho+\xi}(\zeta) \right) \right] \right| \\
\leqslant \quad & \mathbb{E}\left[ \left|\varphi_\ell'(\zeta^{\gamma_F})\,\zeta^{\gamma_F}\right| \left| \frac{1}{B_0(s)} \left\{ \frac{\ell_U(s\zeta) - \ell_U(s)}{a_0(s)} - h_\rho(\zeta) \right\} - h_{\rho+\xi}(\zeta) \right| \right] \\
\leqslant \quad & \varepsilon\, \mathbb{E}\left[ \left|\varphi_\ell'(\zeta^{\gamma_F})\right|\, \zeta^{\gamma_F+\rho+\xi+\delta} \right].
\end{aligned}
$$

Since the expectation in the last line is finite we have that $T_1(s) = o(A_0(s)B_0(s))$. Similar arguments yield that all the expectations in $T_i(s), i = 2, \ldots, 6$, are bounded, from which we deduce that

$$
\begin{aligned}
\mu(s) - \mu(\infty) \quad = \quad & A_0(s)\,\mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\,\zeta^{\gamma_F} h_\rho(\zeta)\right] + A_0(s)\,B_0(s)\,\mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\,\zeta^{\gamma_F}\,h_{\rho+\xi}(\zeta)\right] \\
& + \frac{1}{2} A_0^2(s)\,\mathbb{E}\left[\varphi_\ell''(\zeta^{\gamma_F})\,\zeta^{2\gamma_F} h_\rho^2(\zeta)\right] + o\left(A_0(s)B_0(s)\right) + o\left(A_0^2(s)\right) \\
=: \quad & A_0(s)\,E_1^{(\ell)} + A_0(s)\,B_0(s)\,E_2^{(\ell)} + \frac{1}{2} A_0^2(s)\,E_3^{(\ell)} \\
& + o\left(A_0(s)B_0(s)\right) + o\left(A_0^2(s)\right).
\end{aligned}
$$

Thus, from our decomposition (12)

$$
\begin{aligned}
\mathbb{M}_n^{(\ell)}(k_1) \quad = \quad & M^{(\ell)} + \sigma^{(\ell)}\, \frac{G_{k_1}^{(\ell)}}{\sqrt{k_1}} + A_0\left([1 - F(Z_{n-k_1,n}-)]^{-1}\right)\, E_1^{(\ell)} \\
& + A_0\left([1 - F(Z_{n-k_1,n})]^{-1}\right)\, B_0\left([1 - F(Z_{n-k_1,n})]^{-1}\right)\, E_2^{(\ell)} \\
& + \frac{1}{2}\, A_0^2\left([1 - F(Z_{n-k_1,n})]^{-1}\right)\, E_3^{(\ell)} \\
& + o_{\mathbb{P}}\left(A_0\left([1 - F(Z_{n-k_1,n})]^{-1}\right)\, B_0\left([1 - F(Z_{n-k_1,n})]^{-1}\right)\right) \\
& + o_{\mathbb{P}}\left(A_0^2\left([1 - F(Z_{n-k_1,n})]^{-1}\right)\right),
\end{aligned}
$$

where we have also used that for Pareto-type distributions $(1-F(x-))/(1-F(x)) \to 1$ as $x \to \infty$, along with the regular variation properties of $|A_0|$ and $|B_0|$. From the above, we deduce by a Taylor expansion that for $\tau \in \mathbb{R}^+$ and $\theta \in \mathbb{N}$

$$\begin{aligned}
\left(\frac{\mathbb{M}_n^{(\ell\theta)}(k_1)}{M^{(\ell\theta)}}\right)^{\tau/\theta} &= 1 + \frac{\tau}{\theta}\,\frac{\sigma^{(\ell\theta)}}{M^{(\ell\theta)}}\,\frac{G_{k_1}^{(\ell\theta)}}{\sqrt{k_1}} + \frac{\tau}{\theta}\,\frac{E_1^{(\ell\theta)}}{M^{(\ell\theta)}}\,A_0\left([1-F(Z_{n-k_1,n}-)]^{-1}\right) \\
&\quad + \frac{\tau}{\theta}\,\frac{E_2^{(\ell\theta)}}{M^{(\ell\theta)}}\,A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\,B_0\left([1-F(Z_{n-k_1,n})]^{-1}\right) \\
&\quad + \frac{1}{2}\,\frac{\tau}{\theta}\,\frac{E_3^{(\ell\theta)}}{M^{(\ell\theta)}}\,A_0^2\left([1-F(Z_{n-k_1,n})]^{-1}\right) \\
&\quad + \frac{1}{2}\,\frac{\tau}{\theta}\,\left(\frac{\tau}{\theta}-1\right)\left[\frac{E_1^{(\ell\theta)}}{M^{(\ell\theta)}}\right]^2 A_0^2\left([1-F(Z_{n-k_1,n})]^{-1}\right) \\
&\quad + o_{\mathbb{P}}\left(A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\,B_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\right) \\
&\quad + o_{\mathbb{P}}\left(A_0^2\left([1-F(Z_{n-k_1,n})]^{-1}\right)\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k_1}}\right). \qquad (19)
\end{aligned}$$

Define now

$$\begin{aligned}
d^{(\ell,\theta_1,\theta_2)} &= \frac{E_1^{(\ell\theta_1)}}{\theta_1\,M^{(\ell\theta_1)}} - \frac{E_1^{(\ell\theta_2)}}{\theta_2\,M^{(\ell\theta_2)}}, \\
W_n^{(\ell,\theta_1,\theta_2)} &= \frac{\sigma^{(\ell\theta_1)}}{\theta_1\,M^{(\ell\theta_1)}}\,G_{k_1}^{(\ell\theta_1)} - \frac{\sigma^{(\ell\theta_2)}}{\theta_2\,M^{(\ell\theta_2)}}\,G_{k_1}^{(\ell\theta_2)}, \\
v^{(\ell,\theta_1,\theta_2)} &= \frac{E_2^{(\ell\theta_1)}}{\theta_1\,M^{(\ell\theta_1)}} - \frac{E_2^{(\ell\theta_2)}}{\theta_2\,M^{(\ell\theta_2)}}, \\
u^{(\ell,\theta_1,\theta_2,\tau)} &= \frac{1}{2}\left\{\frac{E_3^{(\ell\theta_1)}}{\theta_1\,M^{(\ell\theta_1)}} - \frac{E_3^{(\ell\theta_2)}}{\theta_2\,M^{(\ell\theta_2)}}\right. \\
&\qquad \left. + (\tau-\theta_1)\left[\frac{E_1^{(\ell\theta_1)}}{\theta_1\,M^{(\ell\theta_1)}}\right]^2 - (\tau-\theta_2)\left[\frac{E_1^{(\ell\theta_2)}}{\theta_2\,M^{(\ell\theta_2)}}\right]^2\right\},
\end{aligned}$$

and introduce

$$D_{k_1,n}^{(\ell,\theta_1,\theta_2,\tau)} := \left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{M^{(\ell\theta_1)}}\right)^{\tau/\theta_1} - \left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{M^{(\ell\theta_2)}}\right)^{\tau/\theta_2}.$$

Clearly from (19), we have

$$\begin{aligned}\frac{D_{k_1,n}^{(\ell,\theta_1,\theta_2,\tau)}}{\tau\, A_0\left([1-F(Z_{n-k_1,n}-)]^{-1}\right)} &= d^{(\ell,\theta_1,\theta_2)} + \frac{1}{\sqrt{k_1}\, A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)} W_n^{(\ell,\theta_1,\theta_2)} \\ &\quad + A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right) u^{(\ell,\theta_1,\theta_2,\tau)} \\ &\quad + B_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\, v^{(\ell,\theta_1,\theta_2)} \\ &\quad + o_{\mathbb{P}}\left(A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\right) \\ &\quad + o_{\mathbb{P}}\left(B_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\right) \\ &\quad + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k_1}\, A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)}\right).\end{aligned}$$

Since $M^{(\ell)} = \gamma_F^{\ell}\,\Gamma(\ell+1)$, we have

$$\begin{aligned}\mathbb{T}_n^{(\tau)}(k_1) &= \frac{D_{k_1,n}^{(1,1,2,\tau)}}{D_{k_1,n}^{(1,2,3,\tau)}} \\ &= \frac{d^{(1,1,2)}}{d^{(1,2,3)}}\left\{1 + \frac{1}{\sqrt{k_1}\, A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)}\left[\frac{W_n^{(1,1,2)}}{d^{(1,1,2)}} - \frac{W_n^{(1,2,3)}}{d^{(1,2,3)}}\right]\right. \\ &\qquad + A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\left[\frac{u^{(1,1,2,\tau)}}{d^{(1,1,2)}} - \frac{u^{(1,2,3,\tau)}}{d^{(1,2,3)}}\right] \\ &\qquad + B_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\left[\frac{v^{(1,1,2)}}{d^{(1,1,2)}} - \frac{v^{(1,2,3)}}{d^{(1,2,3)}}\right] \\ &\qquad + o_{\mathbb{P}}\left(A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\right) + o_{\mathbb{P}}\left(B_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\right) \\ &\qquad \left. + o_{\mathbb{P}}\left(\frac{1}{\sqrt{k_1}\, A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)}\right)\right\}.\end{aligned}$$

We can now deduce that, under our assumptions,

$$
\begin{aligned}
&\sqrt{k_1}\, A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\, \left\{ \mathbb{T}_n^{(\tau)}(k_1) - \frac{d^{(1,1,2)}}{d^{(1,2,3)}} \right\} \\
&\quad = \frac{d^{(1,1,2)}}{d^{(1,2,3)}} \left[ \frac{W_n^{(1,1,2)}}{d^{(1,1,2)}} - \frac{W_n^{(1,2,3)}}{d^{(1,2,3)}} \right] \\
&\qquad + \sqrt{k_1}\, A_0^2\left([1-F(Z_{n-k_1,n})]^{-1}\right)\, \frac{d^{(1,1,2)}}{d^{(1,2,3)}} \left[ \frac{u^{(1,1,2,\tau)}}{d^{(1,1,2)}} - \frac{u^{(1,2,3,\tau)}}{d^{(1,2,3)}} \right] \\
&\qquad + \sqrt{k_1}\, A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)\, B_0\left([1-F(Z_{n-k_1,n})]^{-1}\right) \\
&\qquad\qquad \times \frac{d^{(1,1,2)}}{d^{(1,2,3)}} \left[ \frac{v^{(1,1,2)}}{d^{(1,1,2)}} - \frac{v^{(1,2,3)}}{d^{(1,2,3)}} \right] \\
&\qquad + o_{\mathbb{P}}(1) \\
&\quad = \frac{d^{(1,1,2)}}{d^{(1,2,3)}} \left[ \frac{W_n^{(1,1,2)}}{d^{(1,1,2)}} - \frac{W_n^{(1,2,3)}}{d^{(1,2,3)}} \right] \\
&\qquad + \sqrt{k_1}\, A^2\left([1-F(Z_{n-k_1,n})]^{-1}\right)\, \frac{d^{(1,1,2)}}{d^{(1,2,3)}} \left[ \frac{u^{(1,1,2,\tau)}}{d^{(1,1,2)}} - \frac{u^{(1,2,3,\tau)}}{d^{(1,2,3)}} \right] \\
&\qquad + \sqrt{k_1}\, A\left([1-F(Z_{n-k_1,n})]^{-1}\right)\, B\left([1-F(Z_{n-k_1,n})]^{-1}\right) \\
&\qquad\qquad \times \frac{d^{(1,1,2)}}{d^{(1,2,3)}} \left[ \frac{v^{(1,1,2)}}{d^{(1,1,2)}} - \frac{v^{(1,2,3)}}{d^{(1,2,3)}} \right] \\
&\qquad + o_{\mathbb{P}}(1).
\end{aligned}
$$

Now, since $|A|$, $|B|$, $F$ and $U_H$ are regularly varying, we have, with $\zeta_{n-k_1,n}$ denoting order statistic $(n-k_1)$ in an i.i.d. sample of size $n$ from a strict Pareto(1), that

$$
\begin{aligned}
&\frac{A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)}{A\left([1-F(U_H(n/k_1))]^{-1}\right)} \\
&\qquad \stackrel{d}{=} \frac{A_0\left([1-F(Z_{n-k_1,n})]^{-1}\right)}{A\left([1-F(Z_{n-k_1,n})]^{-1}\right)}\, \frac{A\left([1-F(U_H(\zeta_{n-k_1,n}))]^{-1}\right)}{A\left([1-F(U_H(n/k_1))]^{-1}\right)} \xrightarrow{\mathbb{P}} 1,
\end{aligned}
$$

and, similarly,

$$
\frac{B\left([1-F(Z_{n-k_1,n})]^{-1}\right)}{B\left([1-F(U_H(n/k_1))]^{-1}\right)} \stackrel{d}{=} \frac{B\left([1-F(U_H(\zeta_{n-k_1,n}))]^{-1}\right)}{B\left([1-F(U_H(n/k_1))]^{-1}\right)} \xrightarrow{\mathbb{P}} 1.
$$

Direct computations yield

$$
E_1^{(1)} = \frac{1}{1-\rho}, \qquad E_1^{(2)} = 2\,\gamma_F\,\frac{2-\rho}{(1-\rho)^2}, \qquad E_1^{(3)} = 6\,\gamma_F^2\,\frac{3-3\rho+\rho^2}{(1-\rho)^3},
$$

which, combined with our Lemma 5.1, leads to

$$\sqrt{k_1}\,A\left([1-F(U_H(n/k_1))]^{-1}\right)\left\{\mathbb{T}_n^{(\tau)}(k_1)-\frac{3(1-\rho)}{3-\rho}\right\}\xrightarrow{d}\mathcal{N}\left(m_{\mathbb{T}},\sigma_{\mathbb{T}}^2\right),$$

where

$$\begin{aligned}
m_{\mathbb{T}} &:= \left\{\frac{12(1-\rho)^3}{(1-2\rho)^3(3-\rho^2)}-\frac{3\rho^2(7-20\rho+16\rho^2-4\rho^3)}{(1-2\rho)^2(1-\rho)^2(3-\rho)^2}\frac{1}{\gamma_F}-\frac{\rho(3-2\rho)}{2(1-\rho)^2(3-\rho)}\frac{\tau}{\gamma_F}\right\}\lambda_1\\
&\quad -\frac{6\xi(1-\rho)^3(\rho+\xi)}{\rho(1-\rho-\xi)^3(3-\rho)^2}\,\lambda_2,\\
\sigma_{\mathbb{T}}^2 &:= \frac{36\gamma_G\,\gamma_F^2\,(1-\rho)^6}{\rho^2(3-\rho)^4(\gamma_G-\gamma_F)^5}\,\left\{(2\rho^2-2\rho+1)\gamma_G^4-2\rho(2-\rho)\gamma_G^3\gamma_F\right.\\
&\qquad\qquad \left.+2(\rho^2-2\rho+2)\gamma_G^2\gamma_F^2-2\rho\gamma_G\gamma_F^3+\gamma_F^4\right\}.
\end{aligned}$$

Finally, using the delta-method, Theorem 2.2 follows.

## 5.2 Proof of Theorem 2.3

We follow the lines of proof of Theorem 2.2, but this time only under Model (4), i.e., a second-order framework. In that case, for a fixed $s$, by a Taylor expansion, we have

$$\begin{aligned}
\mu(s)-\mu(\infty) &= \mathbb{E}\left[\varphi_\ell'(u^*(s))\,\{u(s,\zeta)-u(\infty,\zeta)\}\right]\\
&= A_0(s)\,\mathbb{E}\left[\varphi_\ell'(u^*(s))\,\zeta^{\gamma_F}\,h_\rho(\zeta)\right]\\
&\quad +A_0(s)\,\mathbb{E}\left[\varphi_\ell'(u^*(s))\,\left\{\frac{\frac{U_F(s\zeta)}{U_F(s)}-\zeta^{\gamma_F}}{A_0(s)}-\zeta^{\gamma_F}\,h_\rho(\zeta)\right\}\right]
\end{aligned}$$

where $u^*(s)$ is an intermediate random value between $u(s,\zeta)$ and $u(\infty,\zeta)$.
Similarly as for (18), we can show that, as $s\to\infty$,

$$\mathbb{E}\left[\varphi_\ell'(u^*(s))\,\zeta^{\gamma_F}h_\rho(\zeta)\right]\longrightarrow\mathbb{E}\left[\varphi_\ell'(\zeta^{\gamma_F})\,\zeta^{\gamma_F}h_\rho(\zeta)\right].$$

Now, according to Theorem 2.3.9 in [20], for all $\varepsilon,\delta>0$, there exists $s_0=s_0(\varepsilon,\delta)$ such that $\forall s\geqslant s_0$

$$\left|\mathbb{E}\left[\varphi_\ell'(u^*(s))\,\left\{\frac{\frac{U_F(s\zeta)}{U_F(s)}-\zeta^{\gamma_F}}{A_0(s)}-\zeta^{\gamma_F}\,h_\rho(\zeta)\right\}\right]\right|\leqslant\varepsilon\,\mathbb{E}\left[\left|\varphi_\ell'(u^*(s))\right|\,\zeta^{\gamma_F+\rho+\delta}\right],$$

and hence, since the latter expectation is finite, we have, for $s \to \infty$

$$\mathbb{E}\left[\varphi'_\ell(u^*(s))\,\left\{\frac{\frac{U_F(s\zeta)}{U_F(s)}-\zeta^{\gamma_F}}{A_0(s)}-\zeta^{\gamma_F}\,h_\rho(\zeta)\right\}\right]=o(1).$$

Thus, under Model (4), we have

$$\mu(s)-\mu(\infty)=E_1^{(\ell)}\,A(s)+o(A(s)),$$

from which we deduce that

$$\begin{aligned}\mathbb{M}_n^{(\ell)}(k_1) &= M^{(\ell)}+\sigma^{(\ell)}\,\frac{G_{k_1}^{(\ell)}}{\sqrt{k_1}}+A\left([1-F(Z_{n-k_1,n})]^{-1}\right)\,E_1^{(\ell)}\\ &\quad +o_{\mathbb{P}}\left(A\left([1-F(Z_{n-k_1,n})]^{-1}\right)\right).\end{aligned}$$

Now, assuming $\sqrt{k_1}\,A\left([1-F(U_H(n/k_1))]^{-1}\right)\to\infty$ which is implied by Condition (10), we have

$$\frac{\left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{M^{(\ell\theta_1)}}\right)^{\frac{\tau}{\theta_1}}-\left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{M^{(\ell\theta_2)}}\right)^{\frac{\tau}{\theta_2}}}{\tau[1-F(Z_{n-k_1,n})]^{-\rho}}=\left[\frac{\gamma_F}{\theta_1}\,\frac{E_1^{(\ell\theta_1)}}{M^{(\ell\theta_1)}}-\frac{\gamma_F}{\theta_2}\,\frac{E_1^{(\ell\theta_2)}}{M^{(\ell\theta_2)}}\right]\beta+o_{\mathbb{P}}(1),$$

and

$$\frac{\left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{M^{(\ell\theta_1)}}\right)^{\frac{\tau}{\theta_1}\,\kappa}-\left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{M^{(\ell\theta_2)}}\right)^{\frac{\tau}{\theta_2}\,\kappa}}{\tau\,\kappa\,[1-F(Z_{n-k_1,n})]^{-\rho}}=\left[\frac{\gamma_F}{\theta_1}\,\frac{E_1^{(\ell\theta_1)}}{M^{(\ell\theta_1)}}-\frac{\gamma_F}{\theta_2}\,\frac{E_1^{(\ell\theta_2)}}{M^{(\ell\theta_2)}}\right]\beta+o_{\mathbb{P}}(1).$$

This implies that

$$\frac{\left[\frac{\left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{M^{(\ell\theta_1)}}\right)^{\frac{\tau}{\theta_1}}-\left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{M^{(\ell\theta_2)}}\right)^{\frac{\tau}{\theta_2}}}{\tau[1-F(Z_{n-k_1,n})]^{-\rho}}\right]^{\kappa}}{\frac{\left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{M^{(\ell\theta_1)}}\right)^{\frac{\tau}{\theta_1}\,\kappa}-\left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{M^{(\ell\theta_2)}}\right)^{\frac{\tau}{\theta_2}\,\kappa}}{\tau\,\kappa\,[1-F(Z_{n-k_1,n})]^{-\rho}}}=\left[\frac{\gamma_F}{\theta_1}\,\frac{E_1^{(\ell\theta_1)}}{M^{(\ell\theta_1)}}-\frac{\gamma_F}{\theta_2}\,\frac{E_1^{(\ell\theta_2)}}{M^{(\ell\theta_2)}}\right]^{\kappa-1}\beta^{\kappa-1}+o_{\mathbb{P}}(1),$$

which is equivalent to

$$\begin{aligned}&\frac{\left[\frac{\left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{\Gamma(\ell\,\theta_1+1)}\right)^{\frac{\tau}{\theta_1}}-\left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{\Gamma(\ell\,\theta_2+1)}\right)^{\frac{\tau}{\theta_2}}}{\tau[1-F(Z_{n-k_1,n})]^{-\rho}}\right]^{\kappa}}{\frac{\left(\frac{\mathbb{M}_n^{(\ell\theta_1)}(k_1)}{\Gamma(\ell\,\theta_1+1)}\right)^{\frac{\tau}{\theta_1}\,\kappa}-\left(\frac{\mathbb{M}_n^{(\ell\theta_2)}(k_1)}{\Gamma(\ell\,\theta_2+1)}\right)^{\frac{\tau}{\theta_2}\,\kappa}}{\tau\,\kappa\,[1-F(Z_{n-k_1,n})]^{-\rho}}}\\ &=\left[\frac{1}{\theta_1\,\rho}\left\{\left(\frac{1}{1-\rho}\right)^{\ell\,\theta_1}-1\right\}-\frac{1}{\theta_2\,\rho}\left\{\left(\frac{1}{1-\rho}\right)^{\ell\,\theta_2}-1\right\}\right]^{\kappa-1}\beta^{\kappa-1}\\ &\quad +o_{\mathbb{P}}(1).\end{aligned}\tag{20}$$

Using (20) combined with Theorem 2.2 and Condition (10), it is easy to see that

$$\begin{aligned}\frac{\widehat{\beta}^{(c)}(k_1)}{\beta} &= [1-\mathbb{F}_n(Z_{n-k_1,n})]^{\widehat{\rho}^{(c)}(k_1)}[1-F(Z_{n-k_1,n})]^{-\rho}\,(1+o_{\mathbb{P}}(1))\\ &= \left[\frac{1-\mathbb{F}_n(Z_{n-k_1,n})}{1-F(Z_{n-k_1,n})}\right]^{\widehat{\rho}^{(c)}(k_1)}\,[1-F(Z_{n-k_1,n})]^{\widehat{\rho}^{(c)}(k_1)-\rho}\,(1+o_{\mathbb{P}}(1))\\ &= 1+o_{\mathbb{P}}(1),\end{aligned}$$

since, according to [11], we have

$$\left|\frac{1-\mathbb{F}_n(Z_{n-k_1,n})}{1-F(Z_{n-k_1,n})}-1\right| \leqslant \sup_{x\leqslant Z_{n-k_1,n}}\left|\frac{1-\mathbb{F}_n(x)}{1-F(x)}-1\right| = O_{\mathbb{P}}\left(\frac{1}{\sqrt{k_1}}\right), \tag{21}$$

when $F$ is assumed continuous.

### 5.3 Proof of Theorem 2.1.

Using Lemma 5.1 combined with (14) and the assumptions of our theorem, we have the decomposition:

$$\begin{aligned}&\sqrt{k}\left(\widehat{\gamma}_{BC}^{(c)}(k)-\gamma_F\right)\\ &= \sqrt{k}\left(\mathbb{M}_n^{(1)}(k)-\gamma_F\right)-\frac{\mathbb{M}_n^{(1)}(k)\,\widehat{\beta}^{(c)}(k_1)}{1-\widehat{\rho}^{(c)}(k_1)}\sqrt{k}\left\{1-\mathbb{F}_n(Z_{n-k,n})\right\}^{-\widehat{\rho}^{(c)}(k_1)}\\ &= \sigma^{(1)}\,G_k^{(1)}+\frac{1}{1-\rho}\sqrt{k}\,A\left(\left[1-F\left(U_H\left(\frac{n}{k}\right)\right)\right]^{-1}\right)+o_{\mathbb{P}}(1)\\ &\quad -\frac{\mathbb{M}_n^{(1)}(k)\,\widehat{\beta}^{(c)}(k_1)}{1-\widehat{\rho}^{(c)}(k_1)}\sqrt{k}\left\{1-\mathbb{F}_n(Z_{n-k,n})\right\}^{-\widehat{\rho}^{(c)}(k_1)}\\ &= \sigma^{(1)}\,G_k^{(1)}+o_{\mathbb{P}}(1)\\ &\quad +\sqrt{k}\,A\left(\left[1-F\left(U_H\left(\frac{n}{k}\right)\right)\right]^{-1}\right)\left\{\left[\frac{1}{1-\rho}-\frac{1}{1-\widehat{\rho}^{(c)}(k_1)}\right]\right.\\ &\quad -\frac{1}{1-\widehat{\rho}^{(c)}(k_1)}\left[\frac{\mathbb{M}_n^{(1)}(k)}{\gamma_F}\,\frac{\widehat{\beta}^{(c)}(k_1)}{\beta}-1\right]-\frac{1}{1-\widehat{\rho}^{(c)}(k_1)}\,\frac{\mathbb{M}_n^{(1)}(k)}{\gamma_F}\,\frac{\widehat{\beta}^{(c)}(k_1)}{\beta}\\ &\quad \times\left[\left[\frac{1-F(Z_{n-k,n})}{1-F(U_H(n/k))}\right]^{-\widehat{\rho}^{(c)}(k_1)}\left(\frac{1-\mathbb{F}_n(Z_{n-k,n})}{1-F(Z_{n-k,n})}\right)^{-\widehat{\rho}^{(c)}(k_1)}\right.\\ &\qquad\qquad \left.\left.\times\left[1-F\left(U_H\left(\frac{n}{k}\right)\right)\right]^{\rho-\widehat{\rho}^{(c)}(k_1)}-1\right]\right\}\\ &= \sigma^{(1)}\,G_k^{(1)}+o_{\mathbb{P}}(1),\end{aligned}$$

by (21). This achieves the proof of Theorem 2.1.

## 5.4 An auxiliary result and its proof

**Lemma 5.1** *Assume that $F$ and $G$ are of Pareto-type with index $\gamma_F$ and $\gamma_G$, respectively, such that $\gamma_G > \gamma_F$. Under the second-order condition (SOC) and if $\sqrt{k}A\left([1-F\left(U_H\left(n/k\right)\right)]^{-1}\right) \to \lambda \in \mathbb{R}$, we have*

$$\sqrt{k}\left(\mathbb{M}_n^{(1)}(k)-\gamma_F, \mathbb{M}_n^{(2)}(k)-2\,\gamma_F^2, \mathbb{M}_n^{(3)}(k)-6\,\gamma_F^3\right)' \xrightarrow{d} \mathcal{N}\left(\lambda\, b, \Sigma\right),$$

*where*

$$b := \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \text{ and } \Sigma := \begin{pmatrix} \sigma_{1,1} & \sigma_{1,2} & \sigma_{1,3} \\ \sigma_{1,2} & \sigma_{2,2} & \sigma_{2,3} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_{3,3} \end{pmatrix},$$

*with*

$$\begin{aligned}
b_1 &:= \frac{1}{1-\rho}, \\
b_2 &:= \frac{2\,\gamma_F\,(2-\rho)}{(1-\rho)^2}, \\
b_3 &:= \frac{6\,\gamma_F^2\,(3-3\rho+\rho^2)}{(1-\rho)^3}, \\
\sigma_{1,1} &:= \frac{\gamma_G\,\gamma_F^2}{\gamma_G-\gamma_F}, \\
\sigma_{1,2} &:= \frac{2\,\gamma_G\,\gamma_F^3\,(2\,\gamma_G-\gamma_F)}{(\gamma_G-\gamma_F)^2}, \\
\sigma_{1,3} &:= \frac{6\,\gamma_G\,\gamma_F^4\,(3\,\gamma_G^2-3\,\gamma_G\,\gamma_F+\gamma_F^2)}{(\gamma_G-\gamma_F)^3}, \\
\sigma_{2,2} &:= \frac{4\,\gamma_G\,\gamma_F^4\,(5\,\gamma_G^2-4\,\gamma_G\,\gamma_F+\gamma_F^2)}{(\gamma_G-\gamma_F)^3}, \\
\sigma_{2,3} &:= \frac{12\,\gamma_G\,\gamma_F^5\,(9\,\gamma_G^3-10\,\gamma_G^2\,\gamma_F+5\,\gamma_G\,\gamma_F^2-\gamma_F^3)}{(\gamma_G-\gamma_F)^4}, \\
\sigma_{3,3} &:= \frac{36\,\gamma_G\,\gamma_F^6\,(19\,\gamma_G^4-24\,\gamma_G^3\,\gamma_F+16\,\gamma_G^2\,\gamma_F^2-6\,\gamma_G\,\gamma_F^3+\gamma_F^4)}{(\gamma_G-\gamma_F)^5}.
\end{aligned}$$

**Proof of Lemma 5.1.** This lemma is an extension of Lemma E.1 in [7], where we use the fact

that the random variable $W^{\circ}(\varphi_\ell)$, defined in (13), is given for $\ell = 1, 2, 3$ by

$$
\begin{aligned}
W^{\circ}(\log) &= [V^{\circ}]^{\frac{1}{\gamma_G}} \, (\gamma_G - \gamma_F \, \delta^{\circ}) + \gamma_F - \gamma_G, \\
W^{\circ}(\log^2) &= 2 \, [V^{\circ}]^{\frac{1}{\gamma_G}} \, (\log V^{\circ}) \, (\gamma_G - \gamma_F \, \delta^{\circ}) - 2 \, [V^{\circ}]^{\frac{1}{\gamma_G}} \, \left(\gamma_G^2 - \gamma_F \, \gamma_G + \gamma_F^2 \, \delta^{\circ}\right) \\
&\quad +2 \, \left(\gamma_G^2 - \gamma_F \, \gamma_G + \gamma_F^2\right), \\
W^{\circ}(\log^3) &= 3 \, [V^{\circ}]^{\frac{1}{\gamma_G}} \, (\log V^{\circ})^2 \, (\gamma_G - \gamma_F \, \delta^{\circ}) \\
&\quad -6 \, [V^{\circ}]^{\frac{1}{\gamma_G}} \, (\log V^{\circ}) \left(\gamma_G^2 - \gamma_F \, \gamma_G + \gamma_F^2 \, \delta^{\circ}\right) \\
&\quad +6 \, [V^{\circ}]^{\frac{1}{\gamma_G}} \, \left(\gamma_G^3 - \gamma_F \, \gamma_G^2 + \gamma_F^2 \, \gamma_G - \gamma_F^3 \, \delta^{\circ}\right) \\
&\quad -6 \, \left(\gamma_G^3 - \gamma_F \, \gamma_G^2 + \gamma_F^2 \, \gamma_G - \gamma_F^3\right).
\end{aligned}
$$

Tedious computations then achieve the proof of Lemma 5.1.

## Acknowledgement

The research of Martin Bladt was supported by the Carlsberg Foundation, grant CF23-1096. The research of Armelle Guillou was supported by the CNRS grant International Emerging Action IEA-111.

## References

[1] Aban, I.B. and Meerschaert, M.M. (2004). Generalized least-squares estimators for the thickness of heavy tails. *Journal of Statistical Planning and Inference*, 119, 341–352.

[2] Beirlant, J., Bardoutsos, A., de Wet, T. and Gijbels, I. (2016). Bias reduced tail estimation for censored Pareto type distributions. *Statistics & Probability Letters*, 109, 78–88.

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

[4] Beirlant, J., Guillou, A., Dierckx, G. and Fils-Villetard, A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring. *Extremes*, 10, 151–174.

[5] Bladt, M. and Øhlenschlæger, C. (2024). Heterogeneous extremes in the presence of random covariates and censoring. https://arxiv.org/abs/2406.06113.

[6] Bladt, M. and Øhlenschlæger, C. (2024). Censored and extreme losses: functional convergence and applications to tail goodness-of-fit. https://www.arxiv.org/abs/2408.05862.

[7] Bladt M. and Rodionov, I. (2024). Censored extreme value estimation. https://arxiv.org/abs/2312.10499.

[8] Caeiro, F. and Gomes, M.I. (2014). A semi-parametric estimator of a shape second-order parameter. In: Pacheco, A., Santos, R., Oliveira, M., Paulino, C. (eds). *New Advances in Statistical Modeling and Applications. Studies in Theoretical and Applied Statistics*, Springer, 137–144.

[9] Caeiro, F., Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *REVSTAT*, 113–136.

[10] Cheng, S. and Jiang, C. (2001). The Edgeworth expansion for distributions of extreme values. *Science in China*, 44, 427–437.

[11] Csörgő, S. (1996). Universal Gaussian approximations under random censorship. *The Annals of Statistics*, 24, 2744–2778.

[12] Einmahl, J.H.J., Fils-Villetard, A. and Guillou, A. (2008). Statistics of extremes under random censoring. *Bernoulli*, 14, 207–227.

[13] Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *The Annals of Statistics*, 27, 760–781.

[14] 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, 193–213.

[15] Goegebeur, Y., Beirlant, J. and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference*, 140, 2632–2652.

[16] Gomes, M.I., de Haan, L. and Henriques Rodrigues, L. (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.

[17] Gomes, M.I., de Haan, L. and Peng, L. (2002). Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes*, 5, 387–414.

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

[19] Gomes, M.I., Martins, M.J. and Neves, M.M. (2007). Improving second order reduced bias extreme value index estimation. *REVSTAT*, 5, 177–207.

[20] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, New-York, Springer.

[21] Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society: Series B*, 44, 37–42.

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

[23] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3, 1163–1174.

[24] Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53, 457–481.

[25] Peng, L. (1998). Asymptotically unbiased estimators for the extreme-value index. *Statistics & Probability Letters*, 107–115.