



# **Conditional tail moment and reinsurance premium estimation under random right censoring**

Yuri Goegebeur, Armelle Guillou, Jing Qin

## **► To cite this version:**

Yuri Goegebeur, Armelle Guillou, Jing Qin. Conditional tail moment and reinsurance premium estimation under random right censoring. 2023. <hal-03716114v3>

**HAL Id: hal-03716114**

**<https://hal.science/hal-03716114v3>**

Preprint submitted on 5 Jul 2023

**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.



HAL Authorization

# Conditional tail moment and reinsurance premium estimation under random right censoring

Yuri Goegebeur<sup>1†</sup>, Armelle Guillou<sup>2†</sup> and Jing Qin<sup>1†</sup>

<sup>1</sup>Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, Odense, 5230, Denmark.

<sup>2</sup>Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, Strasbourg cedex, 67084, France.

Contributing authors: [Yuri.Goegebeur@imada.sdu.dk](mailto:Yuri.Goegebeur@imada.sdu.dk);  
[armelle.guillou@math.unistra.fr](mailto:armelle.guillou@math.unistra.fr); [qin@imada.sdu.dk](mailto:qin@imada.sdu.dk);

<sup>†</sup>These authors contributed equally to this work.

## Abstract

We propose an estimator of the conditional tail moment (CTM) when the data are subject to random censorship. The variable of main interest and the censoring variable both follow a Pareto-type distribution. We establish the asymptotic properties of our estimator and discuss bias-reduction. Then, the CTM is used to estimate, in case of censorship, the premium principle for excess-of-loss reinsurance. The finite sample properties of the proposed estimators are investigated with a simulation study and we illustrate their practical applicability on a dataset of motor third party liability insurance.

**Keywords:** bias-reduction, conditional tail moment, excess-of-loss reinsurance, Pareto-type distribution, random censorship.

**MSC codes:** 62E20, 62G30, 62G32, 62N01.

# 1 Introduction

The estimation of tail parameters is crucial in many fields where extreme events with possible catastrophic impacts can happen occasionally. In these contexts, Extreme Value Theory (EVT) appears to be the natural tool for modelling this type of events. However, the sample on which the estimation procedure has to be done can be censored. This problem of censoring, where only partial information on a random variable is available, typically that it exceeds a given threshold, is common in many scientific disciplines. For instance, when studying advanced age mortality, one often has that some individuals of a birth cohort are still alive at the time of follow-up, meaning that only a lower bound for their actual lifetime is available. This motivates the study of EVT in the censoring framework, a topic originally considered in the literature by Beirlant et al. (2007) and Einmahl et al. (2008) in all the domains of attraction. However, most recent publications on EVT in the censoring context are concerned with heavy-tailed distributions, motivated by applications in insurance where accurate modelling the upper tail of the claim size distribution is crucial for the risk management but can be difficult since long developments of claims are encountered. This implies that at the evaluation of a portfolio some claims might not be completely settled, and hence they are only known to exceed what has already been paid at the evaluation time. Concretely, this means that the real payments are right censored. We refer for instance to Beirlant et al. (2016, 2018, 2019), Bladt et al. (2021) or Worms and Worms (2014, 2016, 2018), among others.

In this paper, we will also focus on heavy-tailed distributions, i.e., we will assume that our variable of interest  $X$  has a distribution function  $F_X$  which satisfies

$$F_X(x) = 1 - x^{-\frac{1}{\gamma_X}} \ell_X(x), \quad x > 0, \quad (1)$$

where  $\gamma_X > 0$  is called the extreme value index of  $X$ , and  $\ell_X(\cdot)$  is a slowly varying function at infinity, i.e., a function which satisfies

$$\lim_{x \rightarrow \infty} \frac{\ell_X(\lambda x)}{\ell_X(x)} = 1, \text{ for all } \lambda > 0.$$

This variable  $X$  is censored by another random variable, say  $Y$ , independent of  $X$ , which has also a heavy-tailed distribution  $F_Y$  satisfying

$$F_Y(x) = 1 - x^{-\frac{1}{\gamma_Y}} \ell_Y(x), \quad x > 0, \quad (2)$$

with  $\gamma_Y > 0$  and  $\ell_Y(\cdot)$  again a slowly varying function at infinity. In the censoring framework, only  $Z := \min(X, Y)$  is observed together with the indicator  $\delta := \mathbb{1}_{\{X \leq Y\}}$  which specifies whether or not  $X$  has been observed. Given a sample  $(Z_i, \delta_i), 1 \leq i \leq n$ , of independent copies of  $(Z, \delta)$ , our aim

is to make inference on the right tail of the in practice unknown distribution function  $F_X$ . This setup, where both the variable of main interest,  $X$ , and the censoring variable,  $Y$ , are heavy-tailed, and assumed independent is common in the extreme value literature, and was also assumed in, among others, Beirlant et al. (2007), Einmahl et al. (2008) and Bladt et al. (2021). The situation of dependent right censoring in an extreme value context was, to the best of our knowledge, only studied by Stupfler (2019), and this approach requires an assumption on the dependence between  $X$  and  $Y$ .

In particular, first interest is in the estimation of  $\gamma_X$ , despite the fact that the  $X$ -sample was not observed. To understand how to proceed, first remark that, if  $F_Z$  denotes the distribution function of  $Z$ , then (1) and (2) imply that

$$F_Z(x) = 1 - x^{-\frac{1}{\gamma_Z}} \ell_Z(x), \quad x > 0, \quad (3)$$

where  $\gamma_Z := \frac{\gamma_X \gamma_Y}{\gamma_X + \gamma_Y}$ , and  $\ell_Z := \ell_X \ell_Y$ . That means that the distribution function of  $Z$  is also heavy-tailed. Thus, if we turn a blind eye by ignoring the censorship and decide simply to use a classical estimator of the extreme value index such as the Hill estimator (Hill, 1975) based on the  $Z$ -sample and defined as

$$\hat{\gamma}_k^{(H)} := \frac{1}{k} \sum_{i=1}^k \log Z_{n-i+1,n} - \log Z_{n-k,n}, \quad (4)$$

where  $Z_{i,n}$ ,  $1 \leq i \leq n$ , denote the order statistics associated to the  $Z$ -sample and  $k$  the number of random variables taken into account in the estimation procedure, then  $\hat{\gamma}_k^{(H)}$  estimates the extreme value index  $\gamma_Z$  of  $Z$  instead of the one of interest,  $\gamma_X$ , of  $X$ . Since  $\gamma_Z \leq \gamma_X$ , by using  $\hat{\gamma}_k^{(H)}$ , one underestimates the risk if one does not take censoring into consideration. To resolve this problem, Beirlant et al. (2007) proposed to divide the estimator (4) by the proportion of non-censored observations in the  $k$  largest  $Z$ 's, i.e., they introduced the adjusted estimator

$$\hat{\gamma}_k^{(H,c)} := \frac{\frac{1}{k} \sum_{i=1}^k \log Z_{n-i+1,n} - \log Z_{n-k,n}}{\frac{1}{k} \sum_{i=1}^k \delta_{[n-i+1,n]}}, \quad (5)$$

where  $\delta_{[1,n]}, \dots, \delta_{[n,n]}$  denote the  $\delta$ 's corresponding to  $Z_{1,n}, \dots, Z_{n,n}$ , respectively. This denominator in (5) estimates  $\frac{\gamma_Y}{\gamma_X + \gamma_Y}$ , and thus  $\hat{\gamma}_k^{(H,c)}$  is an extreme value index estimator for  $\gamma_X$ .

Einmahl et al. (2008) established, under suitable assumptions, the convergence in distribution of the adjusted estimator  $\hat{\gamma}_k^{(H,c)}$  with the explicit asymptotic bias and variance. Similar results for alternative estimators have been recently proposed in Beirlant et al. (2019) and Bladt et al. (2021).

Since the estimation of extreme value index in the censoring framework is nowadays well-known in the literature, we will assume in the sequel that we have at our disposal an extreme value index estimator adjusted to censoring, denoted  $\hat{\gamma}_k^{(c)}$ , for which the following convergence in distribution holds

$$\sqrt{k} \left( \hat{\gamma}_k^{(c)} - \gamma_X \right) \xrightarrow{d} \Gamma, \quad (6)$$

under suitable conditions. Here  $\Gamma$  follows a normal distribution, with a mean and variance depending on the estimator of the extreme value index used. We refer to Section 2 for specific examples.

Our first aim will be then the estimation of the conditional tail moment (CTM) for a positive random variable  $X$ , defined for some  $\zeta > 0$  as

$$\theta_{p,\zeta} = \mathbb{E} \left( X^\zeta \middle| X > U_X \left( \frac{1}{p} \right) \right),$$

in case  $\mathbb{E}(X^\zeta) < \infty$ , where  $U_X(\cdot) := \inf\{x : F_X(x) \geq 1 - 1/\cdot\} = F_X^\leftarrow(1 - 1/\cdot)$  is the tail quantile function of  $X$ , and when  $X$  is censored. This corresponds, for an insurance company, to the case where some claims can be considered as open in the sense that the company does not know yet the full extent of the loss. Then, using the CTM estimator adapted to censoring, we will look at some well-known risk measures, such as the premium for excess-of-loss reinsurance. By using different values of  $\zeta$ , we will be able in particular to recover the mean and variance of the payment by the reinsurer. Note that this topic has been recently considered in Goegebeur et al. (2022) in case of no-censoring.

The remainder of the paper is organized as follows. In Section 2, we introduce an estimator for the conditional tail moment adapted to censoring and we establish its main asymptotic properties. Then, in Section 3, we look at the premium principle for excess-of-loss reinsurance by providing estimates for the mean and variance of the payment by the reinsurer in case of censoring and we establish their asymptotic behaviors. Finally, Section 4 is devoted to a simulation study, whereas Section 5 illustrates the performance of the estimators on a real dataset from motor third party liability insurance. All the proofs are postponed to the Supplementary Information, which contains also additional simulation results.

## 2 Estimation of the conditional tail moment in case of censoring

We assume that  $X$  and  $Y$  follow a second order Pareto-type model. Let  $RV_\psi$  denote the class of regularly varying functions at infinity with index  $\psi \in \mathbb{R}$ , i.e., positive measurable functions  $f$  satisfying  $f(tx)/f(t) \rightarrow x^\psi$ , as  $t \rightarrow \infty$ ,

for all  $x > 0$ . In the below,  $\bullet$  denotes either  $X$  or  $Y$ .

**Assumption (D)** *The survival function of  $X$  and  $Y$  satisfies*

$$\overline{F}_{\bullet}(x) = A_{\bullet} x^{-\frac{1}{\gamma_{\bullet}}} \left( 1 + \frac{1}{\gamma_{\bullet}} \delta_{\bullet}(x) \right),$$

where  $A_{\bullet} > 0$ ,  $\gamma_{\bullet} > 0$ , and  $|\delta_{\bullet}(\cdot)| \in RV_{-\beta_{\bullet}}$ ,  $\beta_{\bullet} > 0$ .

Clearly, the associated tail quantile function  $U_{\bullet}(\cdot)$  satisfies

$$U_{\bullet}(x) = A_{\bullet}^{\gamma_{\bullet}} x^{\gamma_{\bullet}} (1 + a_{\bullet}(x)) \quad (7)$$

where  $a_{\bullet}(x) = \delta_{\bullet}(U_{\bullet}(x))(1 + o(1))$ , and thus  $|a_{\bullet}(\cdot)| \in RV_{-\beta_{\bullet}\gamma_{\bullet}}$ .

Note that the survival function of  $Z$  also satisfies the form of Assumption (D) with  $A_Z := A_X A_Y$ ,  $\gamma_Z := \frac{\gamma_X \gamma_Y}{\gamma_X + \gamma_Y}$  and  $|\delta_Z|$  being regularly varying with index  $-\beta_Z$ , where  $\beta_Z := \min(\beta_X, \beta_Y)$ , and hence  $U_Z$  will satisfy (7).

We can now expand the CTM under our Pareto-type model on  $X$ .

**Proposition 1** *If  $X$  satisfies Assumption (D) with  $\gamma_X < 1/\zeta$ , then for  $p \downarrow 0$ , we have*

$$\begin{aligned} \theta_{p,\zeta} = & \left[ U_X \left( \frac{1}{p} \right) \right]^{\zeta} \left\{ \frac{1}{1 - \gamma_X \zeta} - \frac{\gamma_X \zeta \beta_X}{(1 - \gamma_X \zeta)(1 - \gamma_X \zeta + \gamma_X \beta_X)} \delta_X \left( U_X \left( \frac{1}{p} \right) \right) \right. \\ & \left. + o \left( \delta_X \left( U_X \left( \frac{1}{p} \right) \right) \right) \right\}. \end{aligned} \quad (8)$$

Note that Proposition 1 is a reformulation of Theorem 2.1 in Goegebeur et al. (2022) under our new Assumption (D) instead of their second order condition.

As a result, a natural estimator for  $\theta_{p,\zeta}$  in the case of censoring is given by

$$\tilde{\theta}_{p,\zeta}^{(c)} := \frac{1}{1 - \hat{\gamma}_k^{(c)} \zeta} \left[ \hat{U}_X^{(c)} \left( \frac{1}{p} \right) \right]^{\zeta}, \quad (9)$$

where  $\hat{U}_X^{(c)} \left( \frac{1}{p} \right)$  is an extreme quantile estimator for  $U_X \left( \frac{1}{p} \right)$  adapted to censoring. For the latter, we can use a Weissman-type estimator (see Weissman, 1978) defined as

$$\hat{U}_X^{(c)} \left( \frac{1}{p} \right) = Z_{n-k,n} \left( \frac{1 - \hat{F}_n(Z_{n-k,n})}{p} \right)^{\hat{\gamma}_k^{(c)}}, \quad (10)$$

where  $\widehat{F}_n(\cdot)$  denotes the Kaplan-Meier product-limit estimator (Kaplan and Meier, 1958), defined as

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

As a preliminary result, we will prove the convergence in distribution of  $\widehat{U}_X^{(c)}\left(\frac{1}{p}\right)$ .

**Theorem 1** Assume  $(\mathcal{D})$  with  $F_X$  continuous and that the convergence (6) holds. Then for  $k, n \rightarrow \infty$ , with  $k/n \rightarrow 0$  and  $\sqrt{k}\delta_X(U_Z(n/k)) \rightarrow \lambda \in \mathbb{R}$ , and  $p$  satisfying  $\frac{\log(k/(np^{(\gamma_X + \gamma_Y)/\gamma_Y}))}{\sqrt{k}} \rightarrow 0$  and  $\frac{k}{np} \rightarrow \infty$ , we have

$$\frac{\sqrt{k}}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{U}_X^{(c)}\left(\frac{1}{p}\right)}{U_X\left(\frac{1}{p}\right)} - 1 \right) \xrightarrow{d} \Gamma.$$

Note that the assumption  $\frac{\log(k/(np^{(\gamma_X + \gamma_Y)/\gamma_Y}))}{\sqrt{k}} \rightarrow 0$  implies that  $\frac{\log((1 - \widehat{F}_n(Z_{n-k,n}))/p)}{\sqrt{k}} \xrightarrow{\mathbb{P}} 0$ , see (B2) in the proof of Theorem 1 in the Supplementary Information.

We have now all the ingredients to prove the convergence in distribution of  $\widehat{\theta}_{p,\zeta}^{(c)}$ .

**Theorem 2** Under the same assumptions as in Theorem 1, for  $\gamma_X < 1/\zeta$ , we have

$$\frac{\sqrt{k}}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{\theta}_{p,\zeta}^{(c)}}{\theta_{p,\zeta}} - 1 \right) \xrightarrow{d} \zeta \Gamma.$$

Note that, in case of no-censoring, this estimator is different from the one proposed in Goegebeur et al. (2022) but has the same limiting distribution under a weaker condition on the tail index, that is  $\gamma_X < 1/\zeta$ , instead of  $\gamma_X < 1/(2\zeta)$ .

If we set  $\widehat{\gamma}_k^{(c)} = \widehat{\gamma}_k^{(H,c)}$  in (9) and (10), then we can make explicit the asymptotic bias and variance of the CTM estimator, denoted in that case by  $\widehat{\theta}_{p,\zeta}^{(H,c)}$ , under the following sub-model of  $(\mathcal{D})$ :

**Assumption  $(\mathcal{D}_H)$**  The survival function of  $X$  and  $Y$  satisfies Assumption  $(\mathcal{D})$  with

$$\delta_{\bullet}(x) = B_{\bullet} x^{-\beta_{\bullet}} (1 + o(1)), B_{\bullet} \in \mathbb{R}, \quad (11)$$

and  $\delta_{\bullet}(x)$  is differentiable, with derivative  $\delta'_{\bullet}(x) = -B_{\bullet}\beta_{\bullet}x^{-\beta_{\bullet}-1}(1 + o(1))$ . Moreover, the function  $x \rightarrow \delta_{\bullet}(x)$  is monotone for  $x$  large enough.

Note that assumption (11) corresponds to the well-known Hall and Welsh (1985) model often used in the literature, for instance in Beirlant et al. (2016).

**Corollary 1** Assume  $(\mathcal{D}_H)$  holds. Then for  $k, n \rightarrow \infty$ , with  $k/n \rightarrow 0$  and  $\sqrt{k}\delta_Z(U_Z(n/k)) \rightarrow \lambda_Z \in \mathbb{R}$ , and  $p$  satisfying  $\frac{\log(k/(np^{(\gamma_X+\gamma_Y)/\gamma_Y}))}{\sqrt{k}} \rightarrow 0$  and  $\frac{k}{np} \rightarrow \infty$ , we have for  $\gamma_X < 1/\zeta$

$$\frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{\theta}_{p,\zeta}^{(H,c)}}{\theta_{p,\zeta}} - 1 \right) \xrightarrow{d} \mathcal{N} \left( \zeta\mu, \zeta^2 \frac{\gamma_X^3}{\gamma_Z} \right),$$

where

$$\mu := \begin{cases} -\frac{\beta_X\gamma_X}{1+\beta_X\gamma_Z} \frac{\gamma_X+\gamma_Y}{\gamma_Y} \lambda_Z & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow \pm\infty, \\ 0 & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow 0, \\ -\frac{\beta_X\gamma_X}{1+\beta_X\gamma_Z} \frac{a(\gamma_X+\gamma_Y)}{\gamma_X+\gamma_Y} \lambda_Z & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow a, \text{ where } |a| \in (0, \infty). \end{cases} \quad (12)$$

Since the estimator for  $\theta_{p,\zeta}$ , after normalisation, inherits the limiting distribution of the estimator that was used for  $\gamma_X$ , we have that the use of  $\widehat{\gamma}_k^{(H,c)}$  will lead to a potentially biased estimator for  $\theta_{p,\zeta}$ . As an alternative, one could use a bias-corrected estimator for  $\gamma_X$ , like the one proposed in Beirlant et al. (2016) when  $\beta_Z$  is known, and given by

$$\widehat{\gamma}_k^{(BC,c)} := \frac{1}{\widehat{d}_k} \left[ \widehat{\gamma}_k^{(H)} + C_{\gamma_Z, \beta_Z} H_{Z,k}^{(-\beta_Z)} \left( H_{Z,k}^{(-\beta_Z)} - \widehat{\gamma}_k^{(H)} \frac{E_{Z,k}^{(c)}(-\beta_Z)}{\widehat{d}_k} \right) \right]$$

where

$$\widehat{d}_k := \frac{1}{k} \sum_{i=1}^k \delta_{[n-i+1, n]},$$

$$C_{\gamma,b} := -\frac{(1+\gamma b)^2(1+2\gamma b)}{\gamma^3 b^2},$$

and, for  $s < 0$ ,

$$H_{Z,k}^{(s)} := \frac{1 - E_{Z,k}(s)}{-s},$$

$$E_{Z,k}(s) := \frac{1}{k} \sum_{i=1}^k \left( \frac{Z_{n-i+1,n}}{Z_{n-k,n}} \right)^s,$$

$$E_{Z,k}^{(c)}(s) := \frac{1}{k} \sum_{i=1}^k \delta_{[n-i+1, n]} \left( \frac{Z_{n-i+1,n}}{Z_{n-k,n}} \right)^s.$$



The limiting distribution of  $\hat{\gamma}_k^{(BC,c)}$ , after proper normalisation, is given in Theorem 1 of Beirlant et al. (2016), and can be used to obtain the following corollary about  $\hat{\theta}_{p,\zeta}^{(c)}$  when using  $\hat{\gamma}_k^{(c)} = \hat{\gamma}_k^{(BC,c)}$ . This bias-corrected estimator for CTM will be denoted as  $\hat{\theta}_{p,\zeta}^{(BC,c)}$ .

**Corollary 2** Assume  $(\mathcal{D}_H)$  holds. Then for  $k, n \rightarrow \infty$ , with  $k/n \rightarrow 0$  and  $\sqrt{k} \left(\frac{n}{k}\right)^{-\min(\gamma_Z \beta_X, \beta_Z)} \rightarrow \lambda \in \mathbb{R}$ , and  $p$  satisfying  $\frac{\log(k/(np^{(\gamma_X + \gamma_Y)/\gamma_Y}))}{\sqrt{k}} \rightarrow 0$  and  $\frac{k}{np} \rightarrow \infty$ , we have for  $\gamma_X < 1/\zeta$

$$\frac{\sqrt{k}}{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\theta}_{p,\zeta}^{(BC,c)}}{\hat{\theta}_{p,\zeta}} - 1 \right) \xrightarrow{d} \mathcal{N} \left( 0, \zeta^2 \frac{\gamma_X^3}{\gamma_Z} \left( \frac{1 + \gamma_Z \beta_Z}{\gamma_Z \beta_Z} \right)^2 \right).$$

Note that the asymptotic variance of the bias-corrected estimator  $\hat{\theta}_{p,\zeta}^{(BC,c)}$  is larger than the one of the uncorrected estimator  $\hat{\theta}_{p,\zeta}^{(H,c)}$ , but this is expected since bias-reduction implies often an increase in the variance. Additionally, remark that the limiting distribution for  $\hat{\gamma}_k^{(BC,c)}$  from Theorem 1 in Beirlant et al. (2016) is derived for the case where  $\beta_Z$  is known. In practice the true  $\beta_Z$  is typically unknown, and for this case Beirlant et al. (2016) propose to replace  $\beta_Z$  by  $\hat{\beta}_Z := -\rho_Z / \hat{\gamma}_k^{(H)}$ , where  $\rho_Z < 0$  is the usual second order parameter in extreme value statistics (in our context  $\rho_Z = -\beta_Z \gamma_Z$ ), and where  $\rho_Z$  is fixed by the user, either at the correct value or mis-specified. In the framework of bias-corrected estimation it is quite common to fix the second order rate parameter at some value, see, e.g., Feuerverger and Hall (1999), Gomes and Martins (2004), Dutang et al. (2014, 2016). Note that in case  $\rho_Z$  is mis-specified then one typically loses theoretically the bias-correction (in the sense that the mean of the limiting normal distribution is no longer zero), but the bias-corrected estimators continue to perform well with respect to bias, and usually outperform estimators that are not corrected for bias; we refer to the simulation results.

### 3 Premium calculation in case of censoring

Our aim in this section is to look at the premium principle for excess-of-loss reinsurance. In particular, we want to estimate in case of censorship and when  $\zeta = 1$  or 2, the quantity

$$\Pi_\zeta(p) := \mathbb{E} \left( (X - U_X(1/p))_+^\zeta \right),$$

where  $A_+ := \max(A, 0)$ . This study is motivated by applications in reinsurance, where  $X$  denotes the claim size,  $(X - U_X(1/p))_+$  the payment by the reinsurer, which arises when the amount of the claim exceeds the retention level  $U_X(1/p)$ . Thus  $\Pi_\zeta(p)$  represents the mean and the second moment of

the payment when  $\zeta = 1$  or  $2$ , respectively. Since some of the claims are still open at the time of the study, i.e., their specific amounts are not yet known, it is crucial to construct estimators which take this censoring into account.

According to Theorem 3.1 in Goegebeur et al. (2022), if  $X$  is of Pareto-type with a continuous distribution function  $F_X$ , we have the following links between the CTMs and  $\Pi_\zeta(p)$  for  $\zeta = 1$  or  $2$ :

$$\begin{aligned}\Pi_1(p) &= p [\theta_{p,1} - U_X(1/p)], \quad \text{if } \gamma_X < 1 \\ \Pi_2(p) &= p [\theta_{p,2} - 2U_X(1/p) \theta_{p,1} + U_X^2(1/p)], \quad \text{if } \gamma_X < 1/2,\end{aligned}$$

from which the natural estimators adapted for censoring follow:

$$\begin{aligned}\hat{\Pi}_1^{(c)}(p) &:= p [\hat{\theta}_{p,1}^{(c)} - \hat{U}_X^{(c)}(1/p)], \quad \text{if } \gamma_X < 1 \\ \hat{\Pi}_2^{(c)}(p) &:= p \left\{ \hat{\theta}_{p,2}^{(c)} - 2\hat{U}_X^{(c)}(1/p) \hat{\theta}_{p,1}^{(c)} + [\hat{U}_X^{(c)}(1/p)]^2 \right\}, \quad \text{if } \gamma_X < 1/2.\end{aligned}$$

From these, we can also introduce an estimator for the variance of the payment by the reinsurer, namely  $\hat{V}^{(c)}(p) := \hat{\Pi}_2^{(c)}(p) - (\hat{\Pi}_1^{(c)}(p))^2$ .

**Theorem 3** *Under the same assumptions as in Theorem 1, we have*

$$\begin{aligned}\frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\Pi}_1^{(c)}(p)}{\Pi_1(p)} - 1 \right) &\xrightarrow{d} \Gamma, \quad \text{if } \gamma_X < 1 \\ \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\Pi}_2^{(c)}(p)}{\Pi_2(p)} - 1 \right) &\xrightarrow{d} 2\Gamma, \quad \text{if } \gamma_X < 1/2.\end{aligned}$$

Again, these estimators of  $\Pi_\zeta(p)$ , for  $\zeta = 1, 2$ , are different from those proposed in Goegebeur et al. (2022) but with similar limiting distributions, and weaker restrictions on  $\gamma_X$ . This is a nice feature since it allows to construct confidence intervals for  $\Pi_\zeta(p)$  for a wide range of values of  $\gamma_X$ .

As before, if we set  $\hat{\gamma}_k^{(c)} = \hat{\gamma}_k^{(H,c)}$  in (9) and (10), then the asymptotic bias and variance of the estimator of  $\Pi_\zeta(p)$ , denoted by  $\hat{\Pi}_\zeta^{(H,c)}(p)$ , can be made explicit.

**Corollary 3** *Under the same assumptions as in Corollary 1, we have for  $\zeta = 1$  or  $2$*

$$\frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\Pi}_\zeta^{(H,c)}(p)}{\Pi_\zeta(p)} - 1 \right) \xrightarrow{d} \mathcal{N} \left( \zeta \mu, \zeta^2 \frac{\gamma_X^3}{\gamma_Z} \right).$$

A similar result can also be obtained for the bias-corrected version, denoted by  $\hat{\Pi}_\zeta^{(BC,c)}(p)$ :

**Corollary 4** Under the same assumptions as in Corollary 2, we have for  $\zeta = 1$  or 2

$$\frac{\sqrt{k}}{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\Pi}_\zeta^{(BC,c)}(p)}{\Pi_\zeta(p)} - 1 \right) \xrightarrow{d} \mathcal{N} \left( 0, \zeta^2 \frac{\gamma_X^3}{\gamma_Z} \left( \frac{1 + \gamma_Z \beta_Z}{\gamma_Z \beta_Z} \right)^2 \right).$$

## 4 Simulation study

In this section we evaluate the finite sample performance of the proposed estimator with a simulation experiment. We simulate from the Burr( $\eta, \lambda, \tau$ ) distribution with distribution function

$$F(x) = 1 - \left( \frac{\eta}{\eta + x^\tau} \right)^\lambda, \quad x > 0,$$

where  $\eta, \lambda, \tau > 0$ . This distribution function satisfies Assumption ( $\mathcal{D}_H$ ) with  $\gamma = 1/(\lambda\tau)$ ,  $\beta = \tau$ . We consider  $X \sim \text{Burr}(1, 1, 4)$ , giving  $\gamma_X = 0.25$  and  $\beta_X = 4$ . We choose the distribution of  $Y$  to obtain a range of asymptotic censoring proportions as follows:

- $Y \sim \text{Burr}(1, 1/4.75, 1)$ , so  $\gamma_Y = 4.75$ ,  $\beta_Y = 1$ , and an asymptotic fraction of censoring of 5%,
- $Y \sim \text{Burr}(1, 1/2.25, 1)$ , so  $\gamma_Y = 2.25$ ,  $\beta_Y = 1$ , and an asymptotic fraction of censoring of 10%,
- $Y \sim \text{Burr}(1, 1, 1)$ , so  $\gamma_Y = 1$ ,  $\beta_Y = 1$ , and an asymptotic fraction of censoring of 20%.
- $Y \sim \text{Burr}(1, 4.2, 1)$ , so  $\gamma_Y = 0.238$ ,  $\beta_Y = 1$ , and an asymptotic fraction of censoring of approximately 51%.

The simulation results for some other distributions and parameter settings can be found in the Supplementary Information.

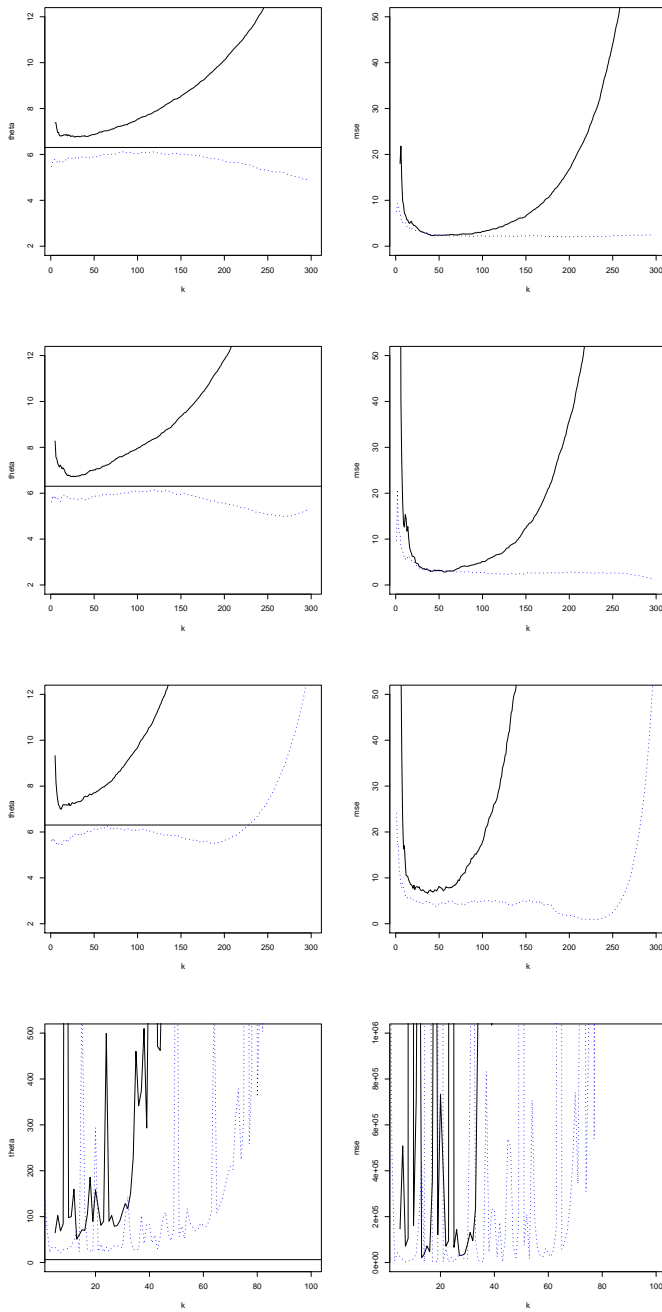
We simulate 200 datasets of size  $n = 500$  from the Burr distributions mentioned above and consider estimation of  $\theta_{1/n,1}$ ,  $\theta_{1/(1.5n),1}$ ,  $\theta_{1/n,2}$ ,  $\Pi_1(1/n)$  and  $V(1/n)$ , except for the distributions leading to 51% of censoring, where we only consider  $\theta_{1/n,1}$  and  $\Pi_1(1/n)$ . In Figure 1 we show the sample mean (left) and sample mean squared error (MSE) (right) of  $\hat{\theta}_{1/n,1}^{(H,c)}$  (black solid line) and  $\hat{\theta}_{1/n,1}^{(BC,c)}$  (blue dashed line), computed over the 200 simulation runs as a function of  $k$  for 5% (first row), 10% (second row), 20% (third row) and 51% (fourth row) of censoring. Figures 2 till 5 are similar, but concern estimation of  $\theta_{1/n,2}$ ,  $\Pi_1(1/n)$ ,  $V(1/n)$  and  $\theta_{1/(1.5n),1}$ , respectively. As for the estimators based on  $\hat{\gamma}_k^{(BC,c)}$ , we follow Beirlant et al. (2016), where it was suggested to replace  $\beta_Z$  by  $\hat{\beta}_Z := -\rho_Z / \hat{\gamma}_k^{(H)}$ , whereafter the estimators for  $\theta_{p,\zeta}$ ,  $\Pi_1(p)$  and  $V(p)$  are computed for several values of  $\rho_Z$ , and plotted as a function of  $k$ . The  $\rho_Z$ -value that leads to the most stable sample paths is then the one to be used for estimation. In our settings  $\rho_Z = -1.5$  gave the most stable results for

all the distributions and estimators under consideration. From the simulations we can draw the following conclusions:

- The estimator for the CTM with  $\zeta = 1$  based on the Hill estimator for  $\gamma$  only works reasonably for small values of  $k$  compared to  $n$ , which is in line with the theoretical results. For increasing values of  $k$  the estimators diverge from the true value and hence show a considerable bias. For the CTM with  $\zeta = 2$  the Hill based estimator performs poorly with a large bias, even at small values of  $k$ .
- On the other hand, the estimator for the CTM based on the bias-corrected estimator of Beirlant et al. (2016) for  $\gamma$  shows an average which is quite stable and close to the true value  $\theta_{p,\zeta}$  for a wide range of values for  $k$ . This confirms again our theoretical results which give that  $\hat{\theta}_{p,\zeta}^{(c)}$ , after normalisation, inherits the limiting distribution of the  $\gamma$ -estimator, so if one uses a bias-corrected estimator for  $\gamma$  then this bias-correction is carried over to  $\hat{\theta}_{p,\zeta}^{(c)}$ . Also,  $\hat{\theta}_{p,\zeta}^{(BC,c)}$  has a minimum value of the MSE which is typically lower than the minimum value of the MSE for  $\hat{\theta}_{p,\zeta}^{(H,c)}$ . Moreover, the bias-corrected estimator has the advantage that the MSE values stay close to the value of the minimum MSE and this for a wide range of values for  $k$ . For practical use of the estimators, the stable sample paths of  $\hat{\theta}_{p,\zeta}^{(BC,c)}$  make the typical issue of choosing  $k$  properly less critical, unlike the situation of  $\hat{\theta}_{p,\zeta}^{(H,c)}$  where proper choice of  $k$  is quite crucial.
- The conclusions above for the CTM also hold for  $\hat{\Pi}_1^{(c)}(1/n)$  and  $\hat{V}^{(c)}(1/n)$ . The estimators for  $\Pi_1(1/n)$  and  $V(1/n)$  based on  $\hat{\gamma}_k^{(BC,c)}$  perform better than those based on  $\hat{\gamma}_k^{(H,c)}$ .
- Estimation of  $\theta_{1/n,2}$  is more challenging than  $\theta_{1/n,1}$ , and similarly, estimation of  $V(1/n)$  is more difficult than  $\Pi_1(1/n)$ .
- As expected, decreasing  $p$  makes the estimation more difficult. This is illustrated in Figure 5 for the estimation of  $\theta_{1/(1.5n),1}$ .
- For all parameters, the estimation gets more challenging as the proportion of censoring increases, as illustrated by MSEs that get elevated with increasing censoring.

Next, we illustrate the selection of  $k$  for  $\hat{\theta}_{p,\zeta}^{(H,c)}$ . In order to make the dependence of  $\hat{\theta}_{p,\zeta}^{(H,c)}$  on  $k$  explicit, we introduce the notation  $\hat{\theta}_{p,\zeta}^{(H,c)}(k)$ . Obviously, small values of  $k$  lead to a large variance of the estimator, while a too large  $k$  leads to a bias issue. Hence, it is natural to determine  $k$  by minimizing an approximation to the asymptotic mean squared error (AMSE) of  $\hat{\theta}_{p,\zeta}^{(H,c)}(k)$ , which, based on Corollary 1, is given by

$$AMSE\left(\hat{\theta}_{p,\zeta}^{(H,c)}(k)\right) = \left(\zeta\theta_{p,\zeta}\log\frac{1-\hat{F}_n(Z_{n-k,n})}{p}\right)^2\left(\frac{\gamma_X^3}{\gamma_Z}\frac{1}{k} + \bar{\mu}_k^2\right),$$



**Fig. 1** Mean (left) and  $MSE$  (right) of  $\hat{\theta}_{1/n,1}^{(H,c)}$  (black solid line) and  $\hat{\theta}_{1/n,1}^{(BC,c)}$  (blue dashed line) as a function of  $k$  for 5% (first row), 10% (second row), 20% (third row) and 51% (fourth row) of censoring.

where

$$\bar{\mu}_k := -\frac{\beta_X \gamma_X}{1 + \beta_X \gamma_Z} \frac{\gamma_X + \gamma_Y}{\gamma_Y} \delta_Z(U_Z(n/k)).$$

The resulting value for  $k$  will be denoted as  $\hat{k}$ . Note that  $\bar{\mu}_k$  is based on the first line of  $\mu$  in (12). This is motivated as follows. The unknown parameters in the above *AMSE* expression clearly need to be estimated in practice, see below. By doing so, it is unlikely that the estimates for  $\beta_X$  and  $\beta_Y$  will be equal, so we ignore the last line for  $\mu$  in (12). The second line in (12) is not useful for determining  $k$ , as the *AMSE* would then only depend on the variance component, leading to a selection of the largest possible value of  $k$ . Whence our proposal to use the first line in (12). Note that this might imply that we incorrectly use the first line of (12) while we should have used the second line. In such situations the value of  $k$  will be chosen too small, though with still an acceptable performance. As an alternative, we could also determine  $k$  by minimising the *AMSE* of  $\hat{\gamma}_k^{(H,c)}$ , given by

$$AMSE(\hat{\gamma}_k^{(H,c)}) = \frac{\gamma_X^3}{\gamma_Z} \frac{1}{k} + \bar{\mu}_k^2,$$

where  $\bar{\mu}_k$  is as above. The resulting value for  $k$  is denoted by  $\tilde{k}$ . As mentioned, the above *AMSE* expressions depend on unknown parameters which need to be estimated from the data. For these we proceed as follows:

- $\gamma_X$  is estimated by  $\hat{\gamma}_k^{(BC,c)}$ ,
- The usual second order parameter  $\rho_X$  in extreme value statistics is given by  $\rho_X = -\beta_X \gamma_X$ . We propose to use the canonical choice  $\rho_X = -1$ , and to estimate  $\beta_X$  as  $1/\hat{\gamma}_k^{(BC,c)}$ ,
- $\gamma_Y/(\gamma_X + \gamma_Y)$  is estimated by the fraction of non-censored observations in the top  $k$  order statistics, i.e., by  $\hat{d}_k$ ,
- $\gamma_Z$  and  $\delta_Z(U_Z(n/k))$  are estimated by fitting the extended Pareto distribution to the relative excesses of the  $Z$ -data, see Beirlant et al. (2009).

In Figure 6 we show the boxplots of  $\hat{\theta}_{1/n,\zeta}^{(H,c)}(\hat{k})/\theta_{1/n,\zeta}$  and  $\hat{\theta}_{1/n,\zeta}^{(H,c)}(\tilde{k})/\theta_{1/n,\zeta}$  for  $\zeta = 1, 2$ , and for the four distributions mentioned above. Overall, and as expected, estimation of the second conditional tail moment is more difficult than the estimation of the first moment, with more bias and variability. Also,  $\hat{\theta}_{1/n,\zeta}^{(H,c)}(\hat{k})$  performs better than  $\hat{\theta}_{1/n,\zeta}^{(H,c)}(\tilde{k})$ , which is also expected as  $\hat{k}$  minimizes (an estimate of)  $AMSE(\hat{\theta}_{p,\zeta}^{(H,c)}(k))$ , while  $\tilde{k}$  minimizes  $AMSE(\hat{\gamma}_k^{(H,c)})$ . Finally, as was indicated by Figure 1, estimation of  $\theta_{1/n,1}$  is very difficult for the dataset with 51% censoring. However, as can be seen from the boxplot in Figure 6, bottom right, combining  $\hat{\theta}_{1/n,1}^{(H,c)}(k)$  with the *AMSE* based procedure for determining  $k$  still leads to acceptable results, though the variability is large.

## 5 Application to insurance data

In this section we analyse a dataset from Motor Third Party Liability Insurance (MTPL) provided by a direct insurance company operating in the EU. The dataset contains the yearly payments of the insurance company, corrected for inflation, for claims by its policyholders over the period 1995–2010. The dataset consists of 837 claims, with about 60% of them being right censored (open) at 2010. We refer to Section 1.3.1 in Albrecher et al. (2017) for more details about this dataset. In Figure 7, left panel, we show the log-claim sizes in chronological order, where the closed claims are depicted in blue and the open claims are in red. As expected, the more recently arrived claims show more censoring than older claims. In order to evaluate the assumption that the underlying distribution of the claim sizes is of Pareto-type we make an adapted Pareto QQ plot in case of right censored data, which is based on the Kaplan-Meier estimator, see Figure 7, right panel. This Pareto QQ plot shows a clear linear pattern in the largest observations which confirms an underlying Pareto-type distribution. We refer to Beirlant et al. (2007) for more details on the construction and interpretation of such QQ plots. In Bladt et al. (2020) it was argued that the assumption of random right censoring is adequate for these data. Figure 8 shows the Hill estimate  $\hat{\gamma}_k^{(H,c)}$  (black line) and the bias-corrected estimate  $\hat{\gamma}_k^{(BC,c)}$  (blue line) for  $\gamma_X$  as a function of  $k$ . The Hill estimate is only stable for the smaller values of  $k$ , say for  $k$  between 50 and 100, while the bias-corrected estimate is stable for  $k$  up to 400. For the values of  $k$  where  $\hat{\gamma}_k^{(H,c)}$  is stable it agrees quite well with  $\hat{\gamma}_k^{(BC,c)}$ . Both the Hill and bias-corrected estimate indicate a value for  $\gamma_X$  above 0.5 but below 1, which in view of the assumption  $\gamma_X < 1/\zeta$  indicates that expected values can be estimated but not second moments (and thus variances). According to these values for  $\gamma$  one can expect estimation results with large variability. The estimation of  $\Pi_1(p)$ , with  $p = 1/n$  as in the simulations, is illustrated in Figure 9 (left panel), where the estimates based on  $\hat{\gamma}_k^{(H,c)}$  are in the black solid line and those based on  $\hat{\gamma}_k^{(BC,c)}$  are in the blue dashed line. The estimates based on  $\hat{\gamma}_k^{(H,c)}$  are very variable and show hardly any stable part, while the premium estimates based on  $\hat{\gamma}_k^{(BC,c)}$  show less variability and are more or less stable for  $k$  in the range 25 till 150. Finally, we illustrate the construction of confidence intervals for  $\Pi_1(p)$ . These will be based on a log-scale version of Theorem 3, as suggested by Drees (2003) in the context of extreme quantile estimation, namely for

$$\frac{\sqrt{k}}{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}} \log \frac{\hat{\Pi}_1^{(c)}(p)}{\Pi_1(p)},$$

which has the same limiting distribution as in Theorem 3. These approximate  $100(1 - \alpha)\%$  confidence intervals are then given by

$$\left[ \hat{\Pi}_1^{(c)}(p) \exp \left( -\Phi^{-1}(1 - \alpha/2) \frac{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}}{\sqrt{k}} \hat{\sigma}_\gamma \right), \right. \\ \left. \hat{\Pi}_1^{(c)}(p) \exp \left( \Phi^{-1}(1 - \alpha/2) \frac{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}}{\sqrt{k}} \hat{\sigma}_\gamma \right) \right],$$

where  $\Phi^{-1}$  denotes the quantile function of the standard normal distribution and  $\hat{\sigma}_\gamma$  is an estimate of the standard deviation of  $\Gamma$ . In Figure 9 (right panel) we show the approximate 95% confidence intervals for  $\Pi_1(p)$  with  $p = 1/n$  as a function of  $k$  in case the estimate is based on  $\hat{\gamma}_k^{(BC,c)}$ . The width of these confidence intervals fluctuates with  $k$  which is due to the variability of the random quantities appearing in the expression for the confidence interval, e.g.,  $\hat{\sigma}_\gamma$ . Note that for this specific dataset the estimation is challenging due to the large values of  $\gamma$ , indicating that the second moment does not exist, as well as the large proportion of censoring.

**Supplementary information.** Proofs of our results can be found in the Supplementary Information, along which additional simulation results.

**Acknowledgments.** The authors sincerely thank the referees, the Associate Editor and Editor for their helpful comments on our paper. This work was supported by the French National Research Agency under Grant ANR-19-CE40-0013-01/ExtremReg project and by the CNRS International Research Network MaDeF.

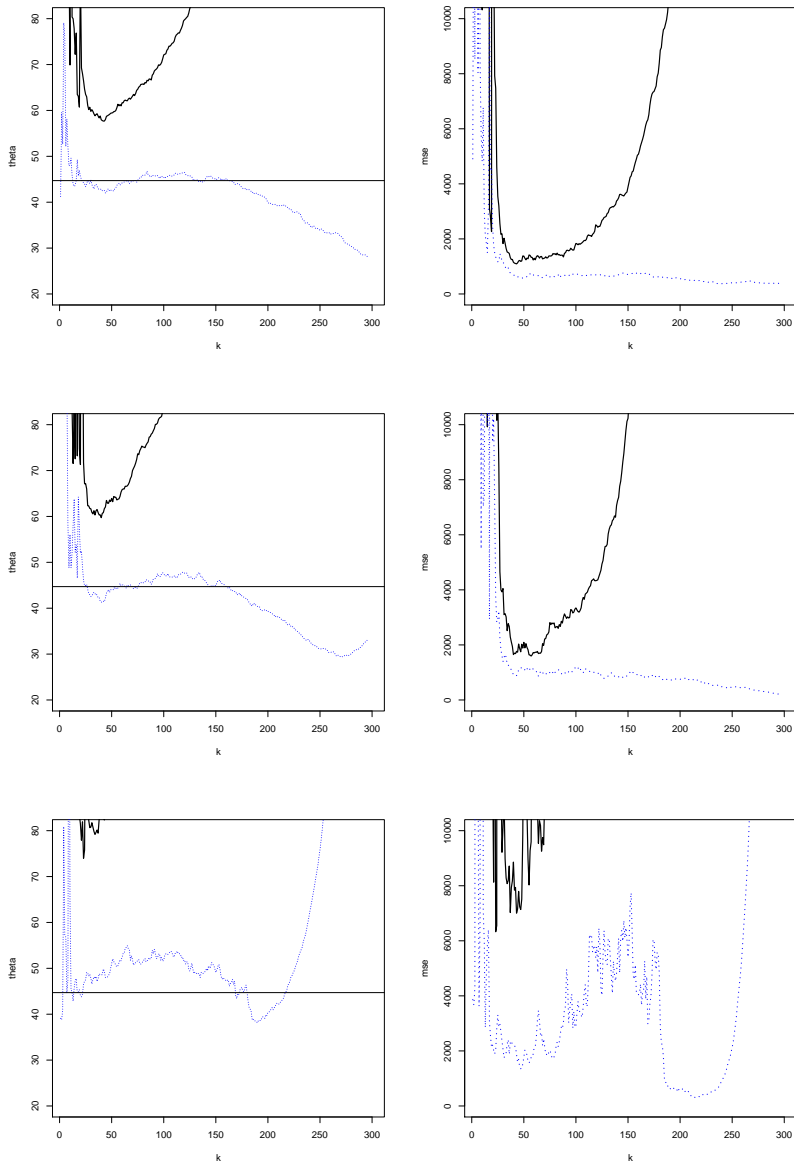
## References

- [1] ALBRECHER, H., BEIRLANT, J., and TEUGELS, J.L. (2017). *Reinsurance: actuarial and statistical aspects*, John Wiley & Sons, Chichester.
- [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., 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.
- [4] BEIRLANT, J., JOOSSENS, E. and SEGERS, J. (2009). Second-order refined peaks-over-threshold modelling for heavy-tailed distributions. *Journal of Statistical Planning and Inference*, **139**, 2800–2815.

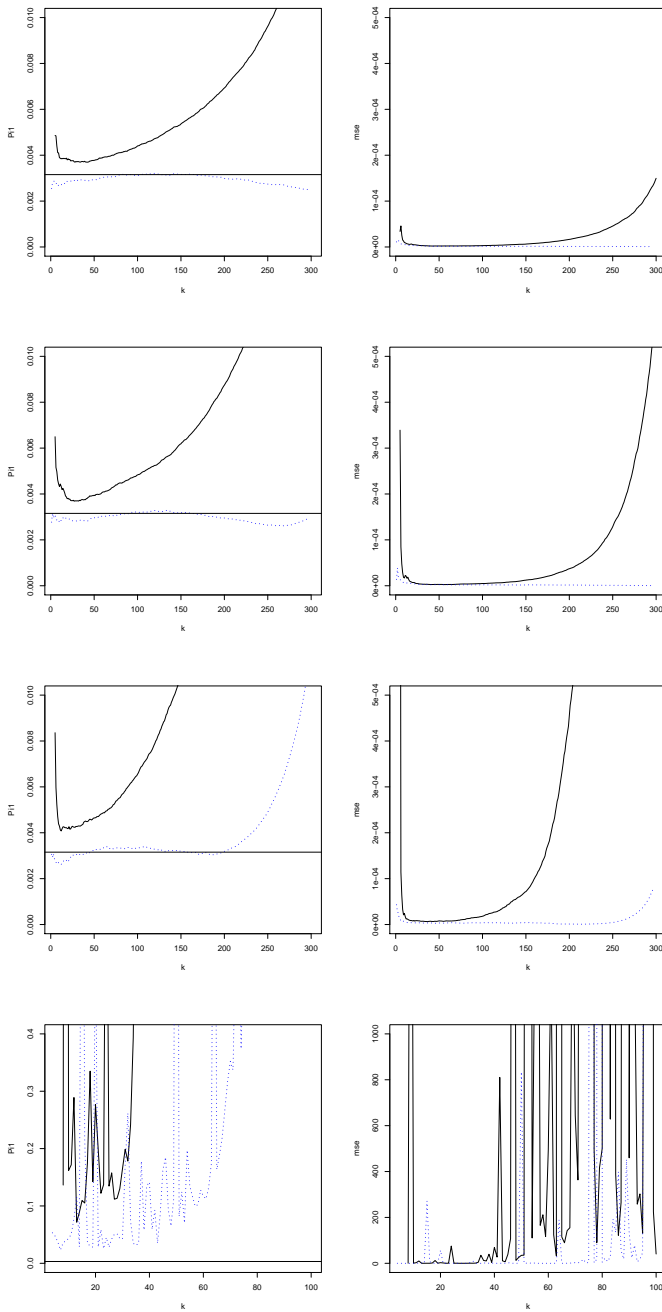


- [5] BEIRLANT, J., MARIBE, G. and VERSTER, A. (2018). Penalized bias reduction in extreme value estimation for censored Pareto-type data, and long-tailed insurance applications. *Insurance: Mathematics and Economics*, **78**, 114–122.
- [6] BEIRLANT, J., WORMS, J. and WORMS, R. (2019). Estimation of the extreme value index in a censorship framework: asymptotic and finite sample behavior. *Journal of Statistical Planning and Inference*, **202**, 31–56.
- [7] BLADT, M., ALBRECHER, H. and BEIRLANT, J. (2020). Combined tail estimation using censored data and expert information. *Scandinavian Actuarial Journal*, **6**, 503–525.
- [8] BLADT, M., ALBRECHER, H. and BEIRLANT, J. (2021). Trimmed extreme value estimators for censored heavy-tailed data. *Electronic Journal of Statistics*, **15**, 3112–3136.
- [9] DREES, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, **9**, 617–657.
- [10] DUTANG, C., GOEGEBEUR, Y. and GUILLOU, A. (2014). Robust and bias-corrected estimation of the coefficient of tail dependence. *Insurance: Mathematics and Economics*, **57**, 46–57.
- [11] DUTANG, C., GOEGEBEUR, Y. and GUILLOU, A. (2016). Robust and bias-corrected estimation of the probability of extreme failure sets. *Sankhya A*, **78**, 52–86.
- [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] GOEGEBEUR, Y., GUILLOU, A., PEDERSEN, T. and QIN, J. (2022). Extreme-value based estimation of the conditional tail moment with application to reinsurance rating. *Insurance: Mathematics and Economics*, **107**, 102–122.
- [15] GOMES, M.I. and MARTINS, M.J. (2004). Bias-reduction and explicit semi-parametric estimation of the tail index. *Journal of Statistical Planning and Inference*, **124**, 361–378.
- [16] HALL, P. and WELSH, A.H. (1985). Adaptive estimates of parameters of regular variation. *The Annals of Statistics*, **13**, 331–341.

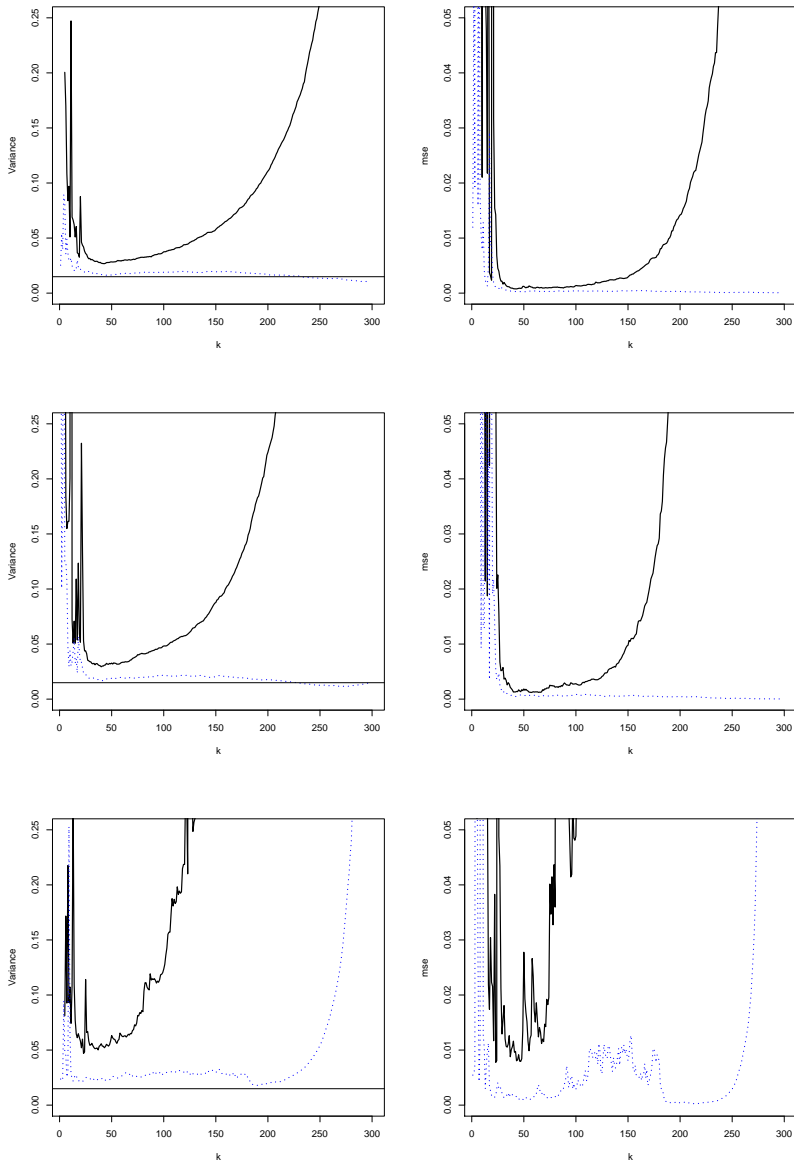
- [17] HILL, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, **3**, 1163–1174.
- [18] KAPLAN, E.L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.
- [19] STUPFLER, G. (2019). On the study of extremes with dependent right-censoring. *Extremes*, **22**, 97–129.
- [20] WEISSMAN, I. (1978). Estimation of parameters and large quantiles based on the  $k$  largest observations. *Journal of the American Statistical Association*, **73**, 812–815.
- [21] WORMS, J. and WORMS, R. (2014). New estimators of the extreme value index under random right censoring, for heavy-tailed distributions. *Extremes*, **17**, 337–358.
- [22] WORMS, J. and WORMS, R. (2016). A Lynden-Bell integral estimator for extremes of randomly truncated data. *Statistics & Probability Letters*, **109**, 106–117.
- [23] WORMS, J. and WORMS, R. (2018). Extreme value statistics for censored data with heavy tails under competing risks. *Metrika*, **81**, 849–889.



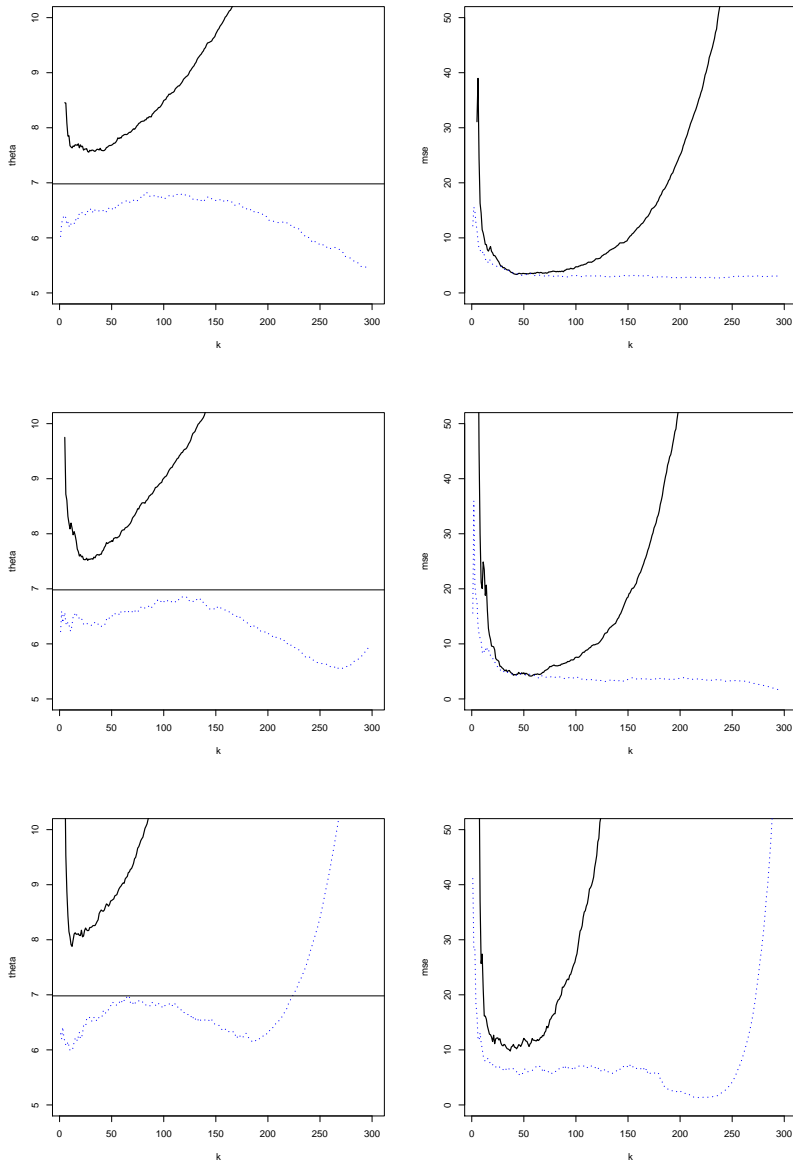
**Fig. 2** Mean (left) and  $MSE$  (right) of  $\hat{\theta}_{1/n,2}^{(H,c)}$  (black solid line) and  $\hat{\theta}_{1/n,2}^{(BC,c)}$  (blue dashed line) as a function of  $k$  for 5% (first row), 10% (second row) and 20% (third row) of censoring.



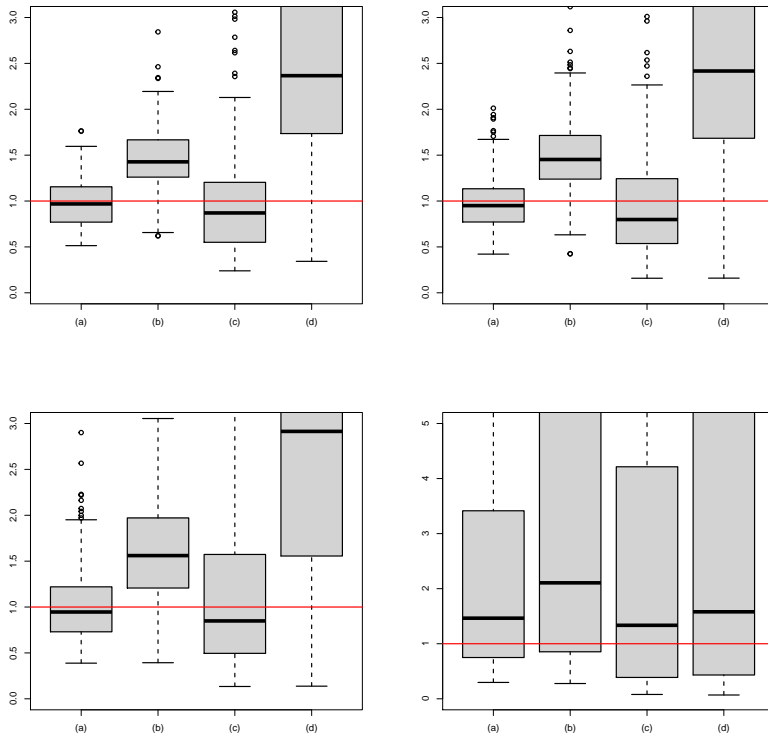
**Fig. 3** Mean (left) and  $MSE$  (right) of  $\hat{\Pi}_1^{(c)}(1/n)$  based on  $\hat{\gamma}_k^{(H,c)}$  (black solid line) and  $\hat{\gamma}_k^{(BC,c)}$  (blue dashed line) as a function of  $k$  for 5% (first row), 10% (second row), 20% (third row) and 51% (fourth row) of censoring.



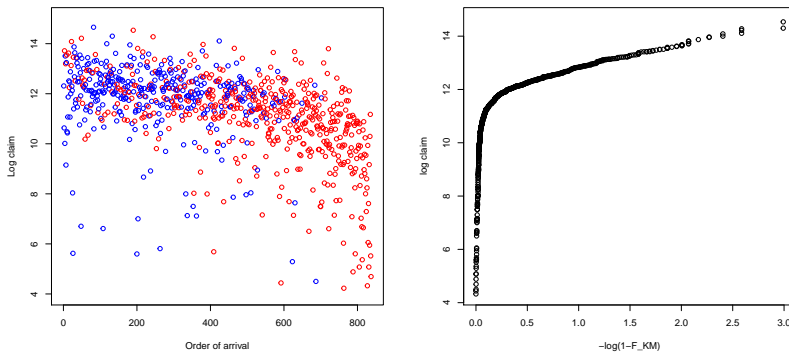
**Fig. 4** Mean (left) and  $MSE$  (right) of  $\widehat{V}^{(c)}(1/n)$  based on  $\widehat{\gamma}_k^{(H,c)}$  (black solid line) and  $\widehat{\gamma}_k^{(BC,c)}$  (blue dashed line) as a function of  $k$  for 5% (first row), 10% (second row) and 20% (third row) of censoring.



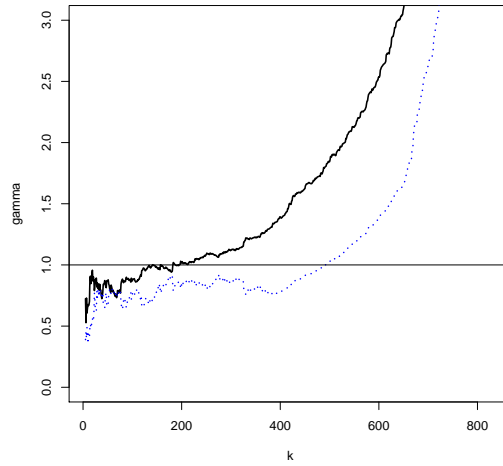
**Fig. 5** Mean (left) and  $MSE$  (right) of  $\hat{\theta}_{1/(1.5n),1}^{(H,c)}$  (black solid line) and  $\hat{\theta}_{1/(1.5n),1}^{(BC,c)}$  (blue dashed line) as a function of  $k$  for 5% (first row), 10% (second row) and 20% (third row) of censoring.



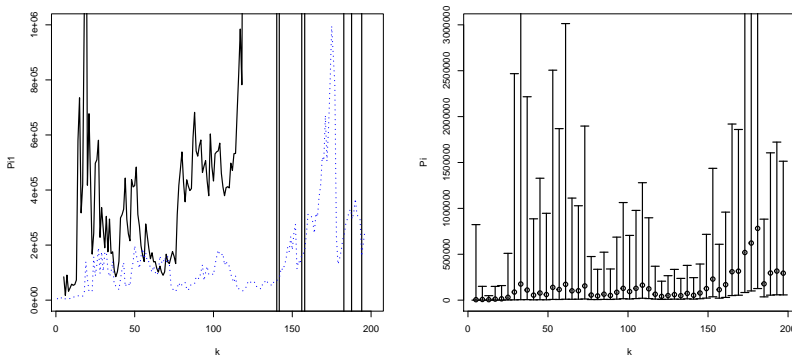
**Fig. 6** Boxplots of (a)  $\hat{\theta}_{1/n,1}^{(H,c)}(\hat{k})/\theta_{1/n,1}$ , (b)  $\hat{\theta}_{1/n,1}^{(H,c)}(\tilde{k})/\theta_{1/n,1}$ , (c)  $\hat{\theta}_{1/n,2}^{(H,c)}(\hat{k})/\theta_{1/n,2}$  and (d)  $\hat{\theta}_{1/n,2}^{(H,c)}(\tilde{k})/\theta_{1/n,2}$  for 5% (top left), 10% (top right), 20% (bottom left) and 51% (bottom right) of censoring.



**Fig. 7** Insurance data. Log-claim sizes in chronological order with closed claims in blue and open claims in red (left), and Pareto QQ plot of the claim sizes (right).



**Fig. 8** Insurance data.  $\hat{\gamma}_k^{(H,c)}$  (black line) and  $\hat{\gamma}_k^{(BC,c)}$  (blue line) as a function of  $k$ .



**Fig. 9** Insurance data.  $\hat{\Pi}_1^{(c)}(1/n)$  based on  $\hat{\gamma}_k^{(H,c)}$  (black solid line) and  $\hat{\gamma}_k^{(BC,c)}$  (blue dashed line) as a function of  $k$  (left) and  $\hat{\Pi}_1^{(c)}(1/n)$  based on  $\hat{\gamma}_k^{(BC,c)}$  with approximate 95% confidence intervals (right).



# Supplementary Information to "Conditional tail moment and reinsurance premium estimation under random right censoring"

Yuri Goegebeur<sup>1†</sup>, Armelle Guillou<sup>2†</sup> and Jing Qin<sup>1†</sup>

<sup>1</sup>Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, Odense, 5230, Denmark.

<sup>2</sup>Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, Strasbourg cedex, 67084, France.

Contributing authors: [Yuri.Goegebeur@imada.sdu.dk](mailto:Yuri.Goegebeur@imada.sdu.dk); [armelle.guillou@math.unistra.fr](mailto:armelle.guillou@math.unistra.fr); [qin@imada.sdu.dk](mailto:qin@imada.sdu.dk);

<sup>†</sup>These authors contributed equally to this work.

## Appendix A Proof of Proposition 1

We adjust the proof of Theorem 2.1 in Goegebeur et al. (2022) with our new Assumption (D). We have

$$\begin{aligned}
 \theta_{p,\zeta} &= -\frac{1}{\bar{F}_X(U_X(1/p))} \int_{U_X(1/p)}^{\infty} z^{\zeta} d\bar{F}_X(z) \\
 &= \left[ U_X\left(\frac{1}{p}\right) \right]^{\zeta} \left\{ 1 + \zeta \int_1^{\infty} z^{\zeta - \frac{1}{\gamma_X} - 1} dz \right. \\
 &\quad + \frac{\zeta}{\gamma_X} \delta_X\left(U_X\left(\frac{1}{p}\right)\right) \int_1^{\infty} z^{\zeta - \frac{1}{\gamma_X} - 1} [z^{-\beta_X} - 1] dz (1 + o(1)) \\
 &\quad \left. + \frac{\zeta}{\gamma_X} \delta_X\left(U_X\left(\frac{1}{p}\right)\right) \int_1^{\infty} z^{\zeta - \frac{1}{\gamma_X} - 1} \left[ \frac{\delta_X(z U_X(1/p))}{\delta_X(U_X(1/p))} - z^{-\beta_X} \right] dz (1 + o(1)) \right\} \\
 &=: \left[ U_X\left(\frac{1}{p}\right) \right]^{\zeta} \left\{ \frac{1}{1 - \gamma_X \zeta} \right.
 \end{aligned}$$

$$-\frac{\gamma_X \zeta \beta_X}{(1 - \gamma_X \zeta)(1 - \gamma_X \zeta + \gamma_X \beta_X)} \delta_X \left( U_X \left( \frac{1}{p} \right) \right) (1 + o(1)) + R_p (1 + o(1)) \Big\},$$

where, according to Proposition B.1.10 in de Haan and Ferreira (2006), for any  $\varepsilon > 0$  and  $\delta \in \left(0, \frac{1}{\gamma_X} - \zeta + \beta_X\right)$ , there exists a  $p_0(\varepsilon, \delta)$  such that for  $p < p_0(\varepsilon, \delta)$  we have

$$\begin{aligned} |R_p| &\leq \frac{\zeta}{\gamma_X} \left| \delta_X \left( U_X \left( \frac{1}{p} \right) \right) \right| \int_1^\infty z^{\zeta - \frac{1}{\gamma_X} - 1} \left| \frac{\delta_X(z U_X(1/p))}{\delta_X(U_X(1/p))} - z^{-\beta_X} \right| dz \\ &\leq \varepsilon \frac{\zeta}{\gamma_X} \left| \delta_X \left( U_X \left( \frac{1}{p} \right) \right) \right| \int_1^\infty z^{\zeta - \frac{1}{\gamma_X} - 1} z^{-\beta_X + \delta} dz \\ &\leq \varepsilon C \left| \delta_X \left( U_X \left( \frac{1}{p} \right) \right) \right|, \end{aligned}$$

for  $C > 0$ . This achieves the proof of Proposition 1.

## Appendix B Proof of Theorem 1

We use the decomposition

$$\begin{aligned} &\frac{\sqrt{k}}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \log \frac{\widehat{U}_X^{(c)} \left( \frac{1}{p} \right)}{U_X \left( \frac{1}{p} \right)} \\ &= \frac{\sqrt{k}}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \log \frac{Z_{n-k,n}}{U_Z(n/k)} \\ &\quad + \sqrt{k} \left( \widehat{\gamma}_k^{(c)} - \gamma_X \right) \\ &\quad + \frac{\sqrt{k}}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \left\{ \log \frac{U_Z(n/k)}{U_X(1/p)} + \gamma_X \log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p} \right\} \\ &= \sqrt{k} \left( \widehat{\gamma}_k^{(c)} - \gamma_X \right) + \frac{1}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \sqrt{k} \log \frac{Z_{n-k,n}}{U_Z(n/k)} \\ &\quad + \frac{\gamma_X}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \sqrt{k} \log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{1 - F_X(Z_{n-k,n})} \\ &\quad + \frac{\sqrt{k}}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \left\{ \log \frac{U_Z(n/k)}{U_X(1/p)} + \gamma_X \log \frac{1 - F_X(Z_{n-k,n})}{p} \right\} \\ &= \sqrt{k} \left( \widehat{\gamma}_k^{(c)} - \gamma_X \right) + \frac{\gamma_X}{\log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{p}} \sqrt{k} \log \frac{1 - \widehat{F}_n(Z_{n-k,n})}{1 - F_X(Z_{n-k,n})} \end{aligned}$$

$$\begin{aligned}
& + \frac{\sqrt{k}}{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}} \left\{ \log \frac{U_Z(n/k)}{U_X(1/p)} + \gamma_X \log \left( \left[ \frac{Z_{n-k,n}}{U_Z(n/k)} \right]^{\frac{1}{\gamma_X}} \frac{1 - F_X(Z_{n-k,n})}{p} \right) \right\} \\
& =: \sqrt{k} \left( \hat{\gamma}_k^{(c)} - \gamma_X \right) + Q_{1,k} + Q_{2,k}.
\end{aligned}$$

Concerning  $Q_{1,k}$ , according to Theorem 2 in Csörgő (1996), if  $F_X$  is continuous, we have

$$\sup_{x \leq Z_{n-k,n}} \left| \frac{\hat{F}_n(x) - F_X(x)}{1 - F_X(x)} \right| = O_{\mathbb{P}} \left( \frac{1}{\sqrt{k}} \right).$$

Thus

$$\begin{aligned}
0 & \leq \frac{1 - \hat{F}_n(Z_{n-k,n})}{1 - F_X(Z_{n-k,n})} \leq 1 + \left| \frac{\hat{F}_n(Z_{n-k,n}) - F_X(Z_{n-k,n})}{1 - F_X(Z_{n-k,n})} \right| \\
& \leq 1 + \sup_{x \leq Z_{n-k,n}} \left| \frac{\hat{F}_n(x) - F_X(x)}{1 - F_X(x)} \right| = 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k}} \right).
\end{aligned}$$

This implies that

$$\sqrt{k} \log \frac{1 - \hat{F}_n(Z_{n-k,n})}{1 - F_X(Z_{n-k,n})} = O_{\mathbb{P}}(1). \quad (\text{B1})$$

Remark that, according to Assumption  $(\mathcal{D})$  and (7), we have

$$\begin{aligned}
\frac{1 - \hat{F}_n(Z_{n-k,n})}{p} &= \frac{1 - F_X(Z_{n-k,n})}{p} \frac{1 - \hat{F}_n(Z_{n-k,n})}{1 - F_X(Z_{n-k,n})} \\
&= \frac{1 - F_X(Z_{n-k,n})}{p} \left( 1 + O_{\mathbb{P}} \left( \frac{1}{\sqrt{k}} \right) \right) \\
&= \frac{A_X}{p} Z_{n-k,n}^{-\frac{1}{\gamma_X}} (1 + o_{\mathbb{P}}(1)) \\
&\stackrel{d}{=} \frac{A_X}{p} (U_Z(\check{Z}_{n-k,n}))^{-\frac{1}{\gamma_X}} (1 + o_{\mathbb{P}}(1)),
\end{aligned}$$

where  $\check{Z}_{n-k,n}$  is order statistic  $n - k$  in a random sample of size  $n$  from the unit Pareto distribution. Since  $\frac{k}{n} \check{Z}_{n-k,n} \xrightarrow{\mathbb{P}} 1$  (see, e.g. Corollary 2.2.2 in de Haan and Ferreira, 2006) and by using the uniform convergence property of regularly varying functions (Theorem B.1.4 in de Haan and Ferreira, 2006) we have that

$$\frac{1 - \hat{F}_n(Z_{n-k,n})}{p} \stackrel{d}{=} \frac{A_X}{p} \left[ U_Z \left( \frac{n}{k} \right) \right]^{-\frac{1}{\gamma_X}} (1 + o_{\mathbb{P}}(1))$$

$$\begin{aligned}
&= [A_X]^{\frac{\gamma_X}{\gamma_X+\gamma_Y}} [A_Y]^{-\frac{\gamma_Y}{\gamma_X+\gamma_Y}} \left(\frac{1}{p}\right)^{\frac{\gamma_X}{\gamma_X+\gamma_Y}} \left(\frac{k}{np}\right)^{\frac{\gamma_Y}{\gamma_X+\gamma_Y}} (1 + o_{\mathbb{P}}(1)) \\
&\xrightarrow{\mathbb{P}} \infty,
\end{aligned} \tag{B2}$$

as soon as  $\frac{k}{n} \rightarrow 0$  and  $\frac{k}{np} \rightarrow \infty$ . Combining (B1) with (B2), we deduce that  $Q_{1,k} = o_{\mathbb{P}}(1)$ .

Now, concerning  $Q_{2,k}$ , remark that, again according to Assumption (D) and (7), we have

$$\begin{aligned}
&\log \frac{U_Z(n/k)}{U_X(1/p)} + \gamma_X \log \left( \left[ \frac{Z_{n-k,n}}{U_Z(n/k)} \right]^{\frac{1}{\gamma_X}} \frac{1 - F_X(Z_{n-k,n})}{p} \right) \\
&= \gamma_X \log \left( 1 + \frac{1}{\gamma_X} \delta_X(Z_{n-k,n}) \right) - \log \left( 1 + a_X \left( \frac{1}{p} \right) \right) \\
&\stackrel{d}{=} \delta_X \left( U_Z \left( \frac{n}{k} \right) \right) (1 + o_{\mathbb{P}}(1)) - \delta_X \left( U_X \left( \frac{1}{p} \right) \right) (1 + o(1)) \\
&= \delta_X \left( U_Z \left( \frac{n}{k} \right) \right) (1 + o_{\mathbb{P}}(1)) + o \left( \delta_X \left( U_X \left( \frac{n}{k} \right) \right) \right), \text{ since } \frac{k}{np} \rightarrow \infty \\
&= \delta_X \left( U_Z \left( \frac{n}{k} \right) \right) (1 + o_{\mathbb{P}}(1)), \text{ since } \gamma_Z \leq \gamma_X.
\end{aligned}$$

Consequently, using again (B2), we deduce that  $Q_{2,k} = o_{\mathbb{P}}(1)$  as soon as  $\sqrt{k} \delta_X \left( U_Z \left( \frac{n}{k} \right) \right) \rightarrow \lambda \in \mathbb{R}$ . This achieves the proof of Theorem 1.

## Appendix C Proof of Theorem 2

Using (6), (8) and (B2), we have the decomposition

$$\begin{aligned}
&\frac{\sqrt{k}}{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\theta}_{p,\zeta}^{(c)}}{\theta_{p,\zeta}} - 1 \right) \\
&= \frac{\sqrt{k}}{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}} \left\{ \left[ \frac{\hat{U}_X^{(c)} \left( \frac{1}{p} \right)}{U_X \left( \frac{1}{p} \right)} \right]^{\zeta} \frac{1 - \gamma_X \zeta}{1 - \hat{\gamma}_k^{(c)} \zeta} \frac{1}{1 + O \left( \delta_X \left( U_X \left( \frac{1}{p} \right) \right) \right)} - 1 \right\} \\
&= \frac{\sqrt{k}}{\log \frac{1 - \hat{F}_n(Z_{n-k,n})}{p}} \left\{ \left[ \frac{\hat{U}_X^{(c)} \left( \frac{1}{p} \right)}{U_X \left( \frac{1}{p} \right)} \right]^{\zeta} \left[ 1 + \frac{\hat{\gamma}_k^{(c)} - \gamma_X}{1 - \hat{\gamma}_k^{(c)} \zeta} \zeta \right] - 1 \right. \\
&\quad \left. + O \left( \delta_X \left( U_X \left( \frac{1}{p} \right) \right) \right) \right\} (1 + o(1))
\end{aligned}$$

$$\begin{aligned}
&= \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \left[ \frac{\widehat{U}_X^{(c)}\left(\frac{1}{p}\right)}{U_X\left(\frac{1}{p}\right)} \right]^\zeta - 1 \right) (1 + o(1)) \\
&\quad + \frac{1}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} O\left(\sqrt{k} \delta_X\left(U_X\left(\frac{1}{p}\right)\right)\right) \\
&\quad + \frac{\zeta}{1 - \widehat{\gamma}_k^{(c)} \zeta} \frac{1}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \sqrt{k} \left(\widehat{\gamma}_k^{(c)} - \gamma_X\right) (1 + o_{\mathbb{P}}(1)) \\
&= \zeta \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{U}_X^{(c)}\left(\frac{1}{p}\right)}{U_X\left(\frac{1}{p}\right)} - 1 \right) + o\left(\sqrt{k} \delta_X\left(U_X\left(\frac{n}{k}\right)\right)\right) + o_{\mathbb{P}}(1) \\
&= \zeta \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{U}_X^{(c)}\left(\frac{1}{p}\right)}{U_X\left(\frac{1}{p}\right)} - 1 \right) + o_{\mathbb{P}}(1). \tag{C3}
\end{aligned}$$

Using Theorem 1, Theorem 2 now follows.

## Appendix D Proof of Corollary 1

We need to use Theorem 1 in Einmahl et al. (2008) which states the convergence in distribution (6) in case  $\widehat{\gamma}_k^{(c)} = \widehat{\gamma}_k^{(H,c)}$ . First, note that under our Assumption ( $\mathcal{D}_H$ ), we have

$$\delta_Z(x) = \begin{cases} \frac{\gamma_Z}{\gamma_X} \delta_X(x) (1 + o(1)) & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow \pm\infty, \\ \frac{\gamma_Z}{\gamma_Y} \delta_Y(x) (1 + o(1)) & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow 0, \\ \frac{\gamma_Z}{\gamma_X} \delta_X(x) \left\{1 + \frac{\gamma_X}{\gamma_Y} \frac{1}{a}\right\} (1 + o(1)) & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow a, \text{ where } |a| \in (0, \infty). \end{cases} \tag{D4}$$

Define now,  $p(z) := \mathbb{P}(\delta = 1 \mid Z = z)$ . Under the assumptions

$$\sqrt{k} \delta_Z\left(U_Z\left(\frac{n}{k}\right)\right) \longrightarrow \lambda_Z \in \mathbb{R}, \tag{D5}$$

$$\frac{1}{\sqrt{k}} \sum_{i=1}^k \left[ p\left(F_Z^{\leftarrow}\left(1 - \frac{i}{n}\right)\right) - \frac{\gamma_Y}{\gamma_X + \gamma_Y} \right] \longrightarrow \alpha \in \mathbb{R}, \tag{D6}$$

and

$$\sqrt{k} \sup_{\{1 - \frac{k}{n} \leq t < 1, |t-s| \leq C \frac{\sqrt{k}}{n}, s < 1\}} \left| p\left(F_Z^{\leftarrow}(t)\right) - p\left(F_Z^{\leftarrow}(s)\right) \right| \longrightarrow 0 \text{ for all } C > 0, \tag{D7}$$

we have

$$\sqrt{k} \left( \hat{\gamma}_k^{(H,c)} - \gamma_X \right) \xrightarrow{d} \mathcal{N} \left( \frac{\gamma_X}{\gamma_Z} [b - \gamma_X \alpha], \sigma^2 \left( \frac{\gamma_X + \gamma_Y}{\gamma_Y} \right)^2 + \frac{\gamma_X^3}{\gamma_Y} \right), \quad (\text{D8})$$

where,  $b := -\frac{\beta_Z \gamma_Z}{1 + \beta_Z \gamma_Z} \lambda_Z$  (respectively  $\sigma^2 := \gamma_Z^2$ ) denotes the bias (respectively, the variance) of the classical Hill estimator  $\hat{\gamma}_k^H$  defined in (4).

Now, direct computations yield

$$\begin{aligned} p(x) - \frac{\gamma_Y}{\gamma_X + \gamma_Y} &= \frac{\bar{F}_Y(x) f_X(x)}{\bar{F}_Y(x) f_X(x) + \bar{F}_X(x) f_Y(x)} - \frac{\gamma_Y}{\gamma_X + \gamma_Y} \\ &= \frac{\gamma_Z}{\gamma_X + \gamma_Y} \{ \beta_X \delta_X(x) - \beta_Y \delta_Y(x) \} + \delta_X(x) o(1) + \delta_Y(x) o(1). \end{aligned}$$

This implies that

$$\begin{aligned} & \frac{1}{\sqrt{k}} \sum_{i=1}^k \left[ p \left( F_Z^{\leftarrow} \left( 1 - \frac{i}{n} \right) \right) - \frac{\gamma_Y}{\gamma_X + \gamma_Y} \right] \\ &= \frac{\gamma_Z \beta_X}{\gamma_X + \gamma_Y} \sqrt{k} \delta_X \left( U_Z \left( \frac{n}{k} \right) \right) \frac{1}{k} \sum_{i=1}^k \frac{\delta_X \left( U_Z \left( \frac{n}{i} \right) \right)}{\delta_X \left( U_Z \left( \frac{n}{k} \right) \right)} (1 + o(1)) \\ & \quad - \frac{\gamma_Z \beta_Y}{\gamma_X + \gamma_Y} \sqrt{k} \delta_Y \left( U_Z \left( \frac{n}{k} \right) \right) \frac{1}{k} \sum_{i=1}^k \frac{\delta_Y \left( U_Z \left( \frac{n}{i} \right) \right)}{\delta_Y \left( U_Z \left( \frac{n}{k} \right) \right)} (1 + o(1)) \\ &= \frac{\gamma_Z \beta_X}{\gamma_X + \gamma_Y} \sqrt{k} \delta_X \left( U_Z \left( \frac{n}{k} \right) \right) \left\{ \frac{1}{k} \sum_{i=1}^k \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_X} \right. \\ & \quad \left. + \frac{1}{k} \sum_{i=1}^k \left[ \frac{\delta_X \left( U_Z \left( \frac{n}{k} \right) \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)}{\delta_X \left( U_Z \left( \frac{n}{k} \right) \right)} - \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_X} \right] \right\} (1 + o(1)) \\ & \quad - \frac{\gamma_Z \beta_Y}{\gamma_X + \gamma_Y} \sqrt{k} \delta_Y \left( U_Z \left( \frac{n}{k} \right) \right) \left\{ \frac{1}{k} \sum_{i=1}^k \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_Y} \right. \\ & \quad \left. + \frac{1}{k} \sum_{i=1}^k \left[ \frac{\delta_Y \left( U_Z \left( \frac{n}{k} \right) \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)}{\delta_Y \left( U_Z \left( \frac{n}{k} \right) \right)} - \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_Y} \right] \right\} (1 + o(1)). \end{aligned}$$

Now, using (7), we have for  $\bullet$  being either  $X$  or  $Y$  and  $n$  large enough:

$$\begin{aligned}
& \frac{1}{k} \sum_{i=1}^k \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_{\bullet}} \\
&= \frac{1}{k} \sum_{i=1}^k \left( \frac{i}{k} \right)^{\beta_{\bullet} \gamma_Z} \left[ \frac{1 + a_Z \left( \frac{n}{i} \right)}{1 + a_Z \left( \frac{n}{k} \right)} \right]^{-\beta_{\bullet}} \\
&= \frac{1}{k} \sum_{i=1}^k \left( \frac{i}{k} \right)^{\beta_{\bullet} \gamma_Z} \left[ 1 + \frac{a_Z \left( \frac{n}{k} \right)}{1 + a_Z \left( \frac{n}{k} \right)} \left\{ \left( \frac{i}{k} \right)^{\beta_Z \gamma_Z} - 1 + \frac{a_Z \left( \frac{n}{i} \right)}{a_Z \left( \frac{n}{k} \right)} - \left( \frac{i}{k} \right)^{\beta_Z \gamma_Z} \right\} \right]^{-\beta_{\bullet}} \\
&= \frac{1}{k} \sum_{i=1}^k \left( \frac{i}{k} \right)^{\beta_{\bullet} \gamma_Z} \left[ 1 + \frac{a_Z \left( \frac{n}{k} \right)}{1 + a_Z \left( \frac{n}{k} \right)} \left\{ \left( \frac{i}{k} \right)^{\beta_Z \gamma_Z} - 1 + o(1) \right\} \right]^{-\beta_{\bullet}},
\end{aligned}$$

by Proposition B.1.10 in de Haan and Ferreira (2006), where the  $o(1)$ -term is uniform in  $i \in \{1, \dots, k\}$ . Thus

$$\begin{aligned}
& \frac{1}{k} \sum_{i=1}^k \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_{\bullet}} \\
&= \frac{1}{k} \sum_{i=1}^k \left( \frac{i}{k} \right)^{\beta_{\bullet} \gamma_Z} - \beta_{\bullet} a_Z \left( \frac{n}{k} \right) \frac{1}{k} \sum_{i=1}^k \left( \frac{i}{k} \right)^{\beta_{\bullet} \gamma_Z} \left\{ \left( \frac{i}{k} \right)^{\beta_Z \gamma_Z} - 1 \right\} + o \left( a_Z \left( \frac{n}{k} \right) \right) \\
&= \frac{1}{1 + \beta_{\bullet} \gamma_Z} + o(1).
\end{aligned}$$

Moreover, using again Proposition B.1.10 in de Haan and Ferreira (2006), we have for  $0 < \delta < \beta_{\bullet}$  and  $n$  large enough

$$\begin{aligned}
& \left| \frac{1}{k} \sum_{i=1}^k \left[ \frac{\delta_{\bullet} \left( U_Z \left( \frac{n}{k} \right) \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)}{\delta_{\bullet} \left( U_Z \left( \frac{n}{k} \right) \right)} - \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_{\bullet}} \right] \right| \\
&\leq \frac{1}{k} \sum_{i=1}^k \left| \frac{\delta_{\bullet} \left( U_Z \left( \frac{n}{k} \right) \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)}{\delta_{\bullet} \left( U_Z \left( \frac{n}{k} \right) \right)} - \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_{\bullet}} \right| \\
&\leq \varepsilon \frac{1}{k} \sum_{i=1}^k \left( \frac{U_Z \left( \frac{n}{i} \right)}{U_Z \left( \frac{n}{k} \right)} \right)^{-\beta_{\bullet} + \delta} \\
&\leq \varepsilon,
\end{aligned}$$

from which we deduce that

$$\begin{aligned}
& \frac{1}{\sqrt{k}} \sum_{i=1}^k \left[ p \left( F_Z^{\leftarrow} \left( 1 - \frac{i}{n} \right) \right) - \frac{\gamma_Y}{\gamma_X + \gamma_Y} \right] \\
&= \frac{\gamma_Z}{\gamma_X + \gamma_Y} \left\{ \frac{\beta_X}{1 + \gamma_Z \beta_X} \sqrt{k} \delta_X \left( U_Z \left( \frac{n}{k} \right) \right) (1 + o(1)) \right. \\
&\quad \left. - \frac{\beta_Y}{1 + \gamma_Z \beta_Y} \sqrt{k} \delta_Y \left( U_Z \left( \frac{n}{k} \right) \right) (1 + o(1)) \right\}.
\end{aligned}$$

Consequently (D6) is satisfied according to (D4), with

$$\alpha = \begin{cases} \frac{\beta_X \gamma_X}{(1 + \beta_X \gamma_Z)(\gamma_X + \gamma_Y)} \lambda_Z & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow \pm\infty, \\ -\frac{\beta_Y \gamma_Y}{(1 + \beta_Y \gamma_Z)(\gamma_X + \gamma_Y)} \lambda_Z & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow 0, \\ \frac{\beta_X \gamma_Z (a-1)}{(1 + \beta_X \gamma_Z)(\gamma_X + \gamma_Y a)} \lambda_Z & \text{if } \frac{\delta_X(x)}{\delta_Y(x)} \rightarrow a, \text{ where } |a| \in (0, \infty). \end{cases} \quad (\text{D9})$$

Concerning (D7), assume first that  $\min(\beta_X, \beta_Y) \gamma_Z > 1$ . Then, using the mean value theorem, for  $s \leq t < 1$  with  $s$  large enough, we have

$$\begin{aligned}
& \left| p \left( F_Z^{\leftarrow}(t) \right) - p \left( F_Z^{\leftarrow}(s) \right) \right| \\
& \leq C_X \left| \delta_X \left( F_Z^{\leftarrow}(t) \right) - \delta_X \left( F_Z^{\leftarrow}(s) \right) \right| + C_Y \left| \delta_Y \left( F_Z^{\leftarrow}(t) \right) - \delta_Y \left( F_Z^{\leftarrow}(s) \right) \right| \\
& \leq C_X (t - s) + C_Y (t - s),
\end{aligned}$$

where  $C_X, C_Y > 0$  (possibly different from line to line). Thus (D7) is clearly satisfied.

Assume now that  $\max(\beta_X, \beta_Y) \gamma_Z < 1$ . Then, for  $s \leq t < 1$  with  $s$  large enough, using the monotonicity of  $\delta_{\bullet}(\cdot)$  for  $\bullet$  either  $X$  or  $Y$ , we have

$$\begin{aligned}
& \sqrt{k} \sup_{\{1 - \frac{k}{n} \leq t < 1, |t-s| \leq C \frac{\sqrt{k}}{n}, s < 1\}} \left| p \left( F_Z^{\leftarrow}(t) \right) - p \left( F_Z^{\leftarrow}(s) \right) \right| \\
& \leq \text{sgn}(B_X) C_X \sqrt{k} \sup_{\{1 - \frac{k}{n} \leq t < 1, t-s \leq C \frac{\sqrt{k}}{n}, s < 1\}} \left[ \delta_X \left( F_Z^{\leftarrow}(s) \right) - \delta_X \left( F_Z^{\leftarrow}(t) \right) \right] \\
& \quad + \text{sgn}(B_Y) C_Y \sqrt{k} \sup_{\{1 - \frac{k}{n} \leq t < 1, t-s \leq C \frac{\sqrt{k}}{n}, s < 1\}} \left[ \delta_Y \left( F_Z^{\leftarrow}(s) \right) - \delta_Y \left( F_Z^{\leftarrow}(t) \right) \right] \\
& \leq \text{sgn}(B_X) C_X \sqrt{k} \sup_{\{1 - \frac{k}{n} \leq t < 1\}} \left[ \delta_X \left( F_Z^{\leftarrow} \left( t - C \frac{\sqrt{k}}{n} \right) \right) - \delta_X \left( F_Z^{\leftarrow}(t) \right) \right] \\
& \quad + \text{sgn}(B_Y) C_Y \sqrt{k} \sup_{\{1 - \frac{k}{n} \leq t < 1\}} \left[ \delta_Y \left( F_Z^{\leftarrow} \left( t - C \frac{\sqrt{k}}{n} \right) \right) - \delta_Y \left( F_Z^{\leftarrow}(t) \right) \right] \\
& = o \left( \sqrt{k} \delta_X \left( U_Z \left( \frac{n}{k} \right) \right) \right) + o \left( \sqrt{k} \delta_Y \left( U_Z \left( \frac{n}{k} \right) \right) \right) \\
& = o \left( \sqrt{k} \delta_Z \left( U_Z \left( \frac{n}{k} \right) \right) \right)
\end{aligned}$$



$$= o(1),$$

by assumption. In case  $\min(\beta_X, \beta_Y)\gamma_Z < 1 < \max(\beta_X, \beta_Y)\gamma_Z$ , we can combine the two above ideas of proof to conclude that the convergence (D7) also holds. Combining (D8) with (D9), Corollary 1 holds.

## Appendix E Proof of Theorem 3

We have the decomposition

$$\begin{aligned} & \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\Pi}_1^{(c)}(p)}{\Pi_1(p)} - 1 \right) \\ &= \frac{\theta_{p,1}}{\theta_{p,1} - U_X(1/p)} \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\theta}_{p,1}^{(c)}}{\theta_{p,1}} - 1 \right) \\ & \quad - \frac{U_X(1/p)}{\theta_{p,1} - U_X(1/p)} \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{U}_X^{(c)}(1/p)}{U_X(1/p)} - 1 \right) \\ &= \frac{1}{\gamma_X} \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\theta}_{p,1}^{(c)}}{\theta_{p,1}} - 1 \right) \\ & \quad - \frac{1-\gamma_X}{\gamma_X} \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{U}_X^{(c)}(1/p)}{U_X(1/p)} - 1 \right) + o_{\mathbb{P}}(1) \end{aligned}$$

by Proposition 1. Now, using (C3), we have

$$\frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\Pi}_1^{(c)}(p)}{\Pi_1(p)} - 1 \right) = \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{U}_X^{(c)}(1/p)}{U_X(1/p)} - 1 \right) + o_{\mathbb{P}}(1),$$

from which the first part of Theorem 3 follows.

The second part of Theorem 3 can be proved similarly since by Proposition 1 and (C3), we have

$$\begin{aligned} & \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\Pi}_2^{(c)}(p)}{\Pi_2(p)} - 1 \right) \\ &= \frac{\theta_{p,2}}{\theta_{p,2} - 2U_X(1/p)\theta_{p,1} + U_X^2(1/p)} \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{\theta}_{p,2}^{(c)}}{\theta_{p,2}} - 1 \right) \\ & \quad - \frac{2\hat{\theta}_{p,1}^{(c)}U_X(1/p)}{\theta_{p,2} - 2U_X(1/p)\theta_{p,1} + U_X^2(1/p)} \frac{\sqrt{k}}{\log \frac{1-\hat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\hat{U}_X^{(c)}(1/p)}{U_X(1/p)} - 1 \right) \end{aligned}$$

$$\begin{aligned}
& - \frac{2\theta_{p,1} U_X(1/p)}{\theta_{p,2} - 2U_X(1/p)\theta_{p,1} + U_X^2(1/p)} \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{\theta}_{p,1}^{(c)}}{\theta_{p,1}} - 1 \right) \\
& + \frac{U_X^2(1/p)}{\theta_{p,2} - 2U_X(1/p)\theta_{p,1} + U_X^2(1/p)} \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \left[ \frac{\widehat{U}_X^{(c)}(1/p)}{U_X(1/p)} \right]^2 - 1 \right) \\
& = \frac{1-\gamma_X}{2\gamma_X^2} \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{\theta}_{p,2}^{(c)}}{\theta_{p,2}} - 1 \right) \\
& - \frac{1-2\gamma_X}{\gamma_X^2} \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{U}_X^{(c)}(1/p)}{U_X(1/p)} - 1 \right) \\
& - \frac{1-2\gamma_X}{\gamma_X^2} \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{\theta}_{p,1}^{(c)}}{\theta_{p,1}} - 1 \right) \\
& + \frac{(1-\gamma_X)(1-2\gamma_X)}{2\gamma_X^2} \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \left[ \frac{\widehat{U}_X^{(c)}(1/p)}{U_X(1/p)} \right]^2 - 1 \right) + o_{\mathbb{P}}(1) \\
& = 2 \frac{\sqrt{k}}{\log \frac{1-\widehat{F}_n(Z_{n-k,n})}{p}} \left( \frac{\widehat{U}_X^{(c)}(1/p)}{U_X(1/p)} - 1 \right) + o_{\mathbb{P}}(1),
\end{aligned}$$

which ends the proof of Theorem 3.

## Appendix F Additional simulation results

In this section we report the results of some additional simulation experiments. Besides the Burr distribution, already introduced in the main paper, we also consider the following models:

- The Fréchet distribution,  $\text{Fréchet}(\alpha)$ ,  $\alpha > 0$ , with distribution function

$$F(x) = e^{-x^{-\alpha}}, \quad x > 0.$$

For this model  $\gamma = 1/\alpha$  and  $\beta = \alpha$ .

- The generalized Pareto distribution,  $\text{GPD}(\gamma, \sigma)$ ,  $\gamma, \sigma > 0$ , with distribution function

$$F(x) = 1 - \left( 1 + \frac{\gamma}{\sigma} x \right)^{-1/\gamma}, \quad x > 0.$$

For this model  $\beta = 1$ .

In particular, we consider the following setups, leading to 10% censoring:

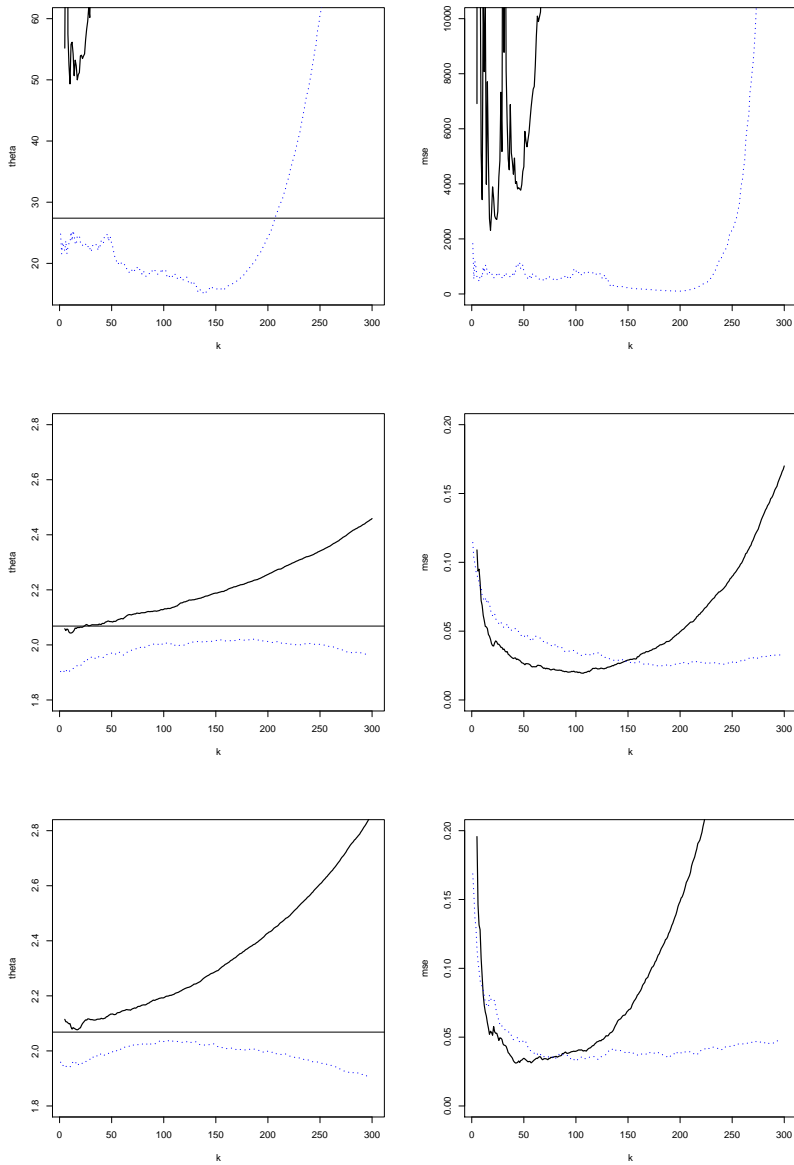
1.  $X \sim \text{GPD}(0.3, 1)$  giving  $\gamma_X = 0.3$ ,  $\beta_X = 1$ , and  $Y \sim \text{Burr}(1, 1/5.4, 2)$  leading to  $\beta_Y = 2$ .

2.  $X \sim \text{Fréchet}(10)$  giving  $\gamma_X = 0.1$ ,  $\beta_X = 10$ , and  $Y \sim \text{Burr}(1, 1/4.5, 5)$  leading to  $\beta_Y = 5$ .
3.  $X \sim \text{Burr}(1, 1, 10)$  giving  $\gamma_X = 0.1$ ,  $\beta_X = 10$ , and  $Y \sim \text{Burr}(1, 1/9, 10)$  leading to  $\beta_Y = 10$ .

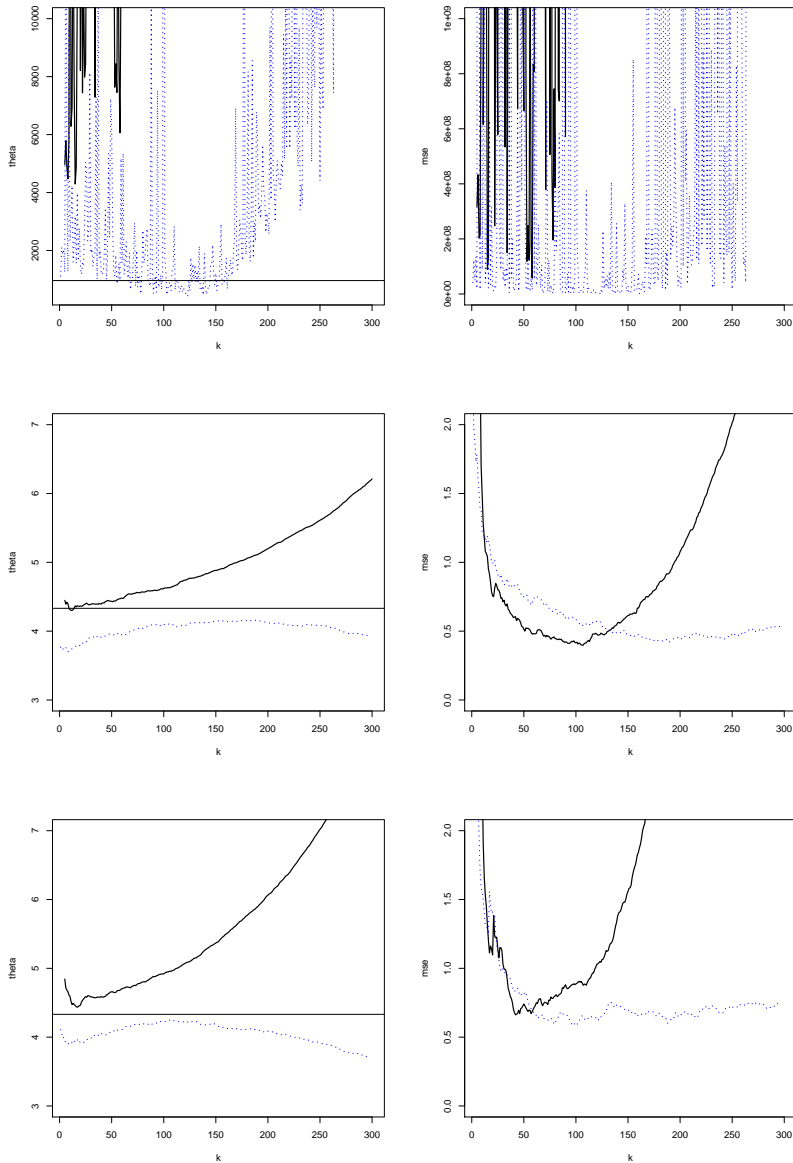
Overall, the same conclusions can be drawn as those mentioned in the simulation section of the main paper. Note that Setting 1 is very difficult as the GPD has second order parameter  $\rho_X = -\gamma_X$ , so in our case  $\rho_X = -0.3$ . It is well-known in extreme value statistics that  $\rho_X$  close to zero is very challenging as the convergence rate in the second order condition is then slow, which leads typically to more biased estimators. Secondly,  $\gamma_X = 0.3$  is the largest value considered in our simulation setup, and clearly makes the estimation of second moments challenging for the given size of the dataset.

## References

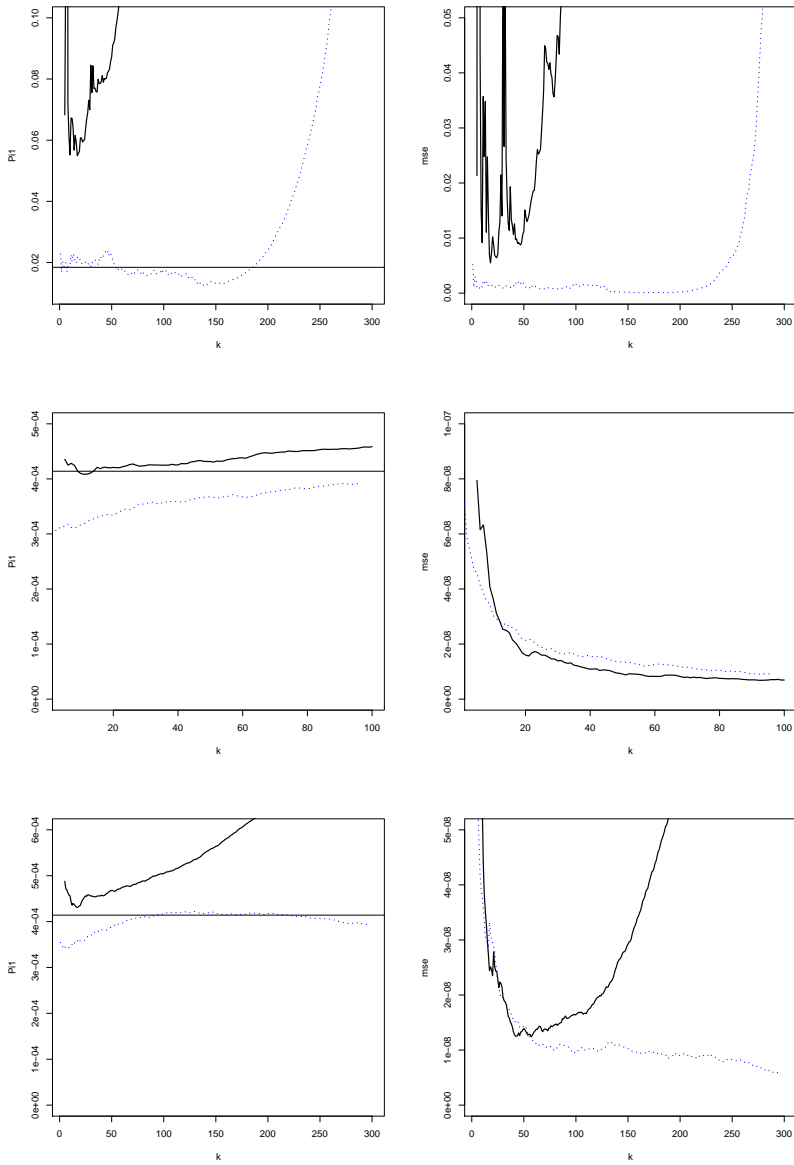
- [1] CSÖRGŐ, S. (1996). Universal Gaussian approximations under random censorship. *The Annals of Statistics*, **24**, 2744–2778.
- [2] EINMAHL, J. H. J., FILS-VILLETARD, A. and GUILLOU, A. (2008). Statistics of extremes under random censoring. *Bernoulli*, **14**, 207–227.
- [3] GOEGEBEUR, Y., GUILLOU, A., PEDERSEN, T. and QIN, J. (2022). Extreme-value based estimation of the conditional tail moment with application to reinsurance rating. *Insurance: Mathematics and Economics*, **107**, 102–122.
- [4] DE HAAN, L. and FERREIRA, A. (2006). *Extreme value theory - An introduction*, Springer.



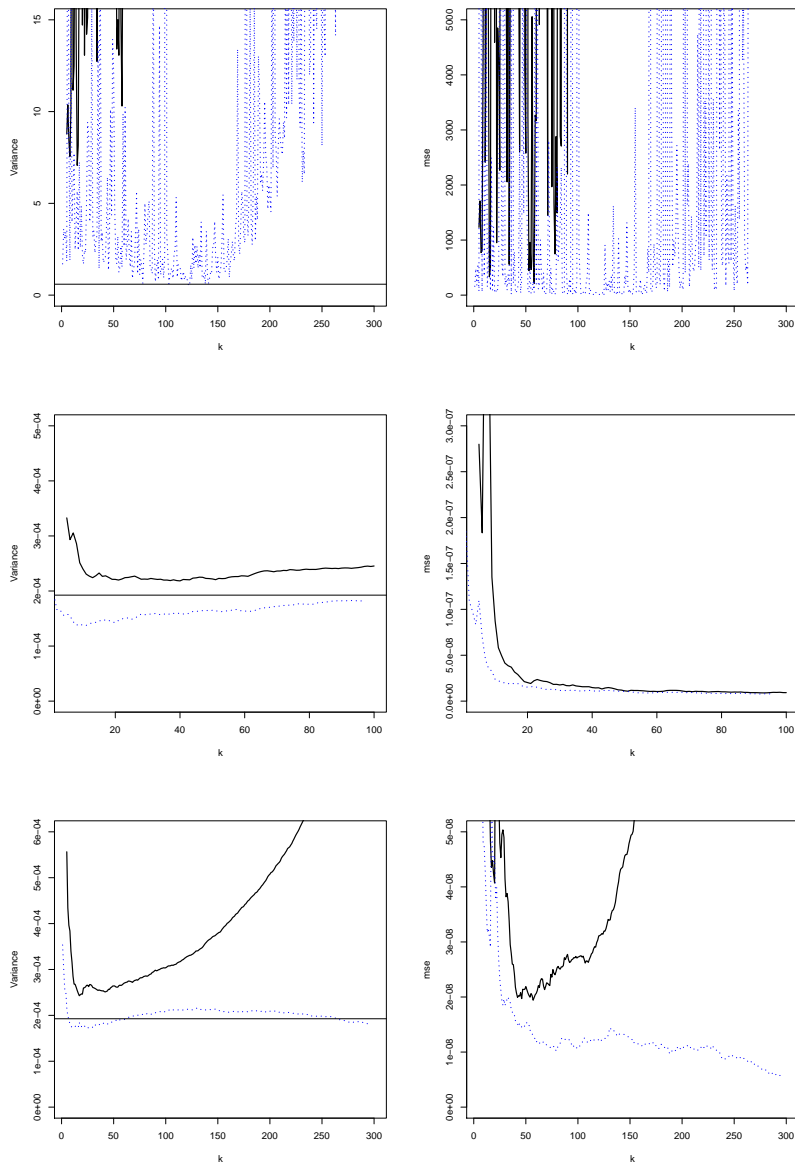
**Fig. F1** Mean (left) and  $MSE$  (right) of  $\hat{\theta}_{1/n,1}^{(H,c)}$  (black solid line) and  $\hat{\theta}_{1/n,1}^{(BC,c)}$  (blue dashed line) as a function of  $k$  for Setup 1 (first row), Setup 2 (second row) and Setup 3 (third row).



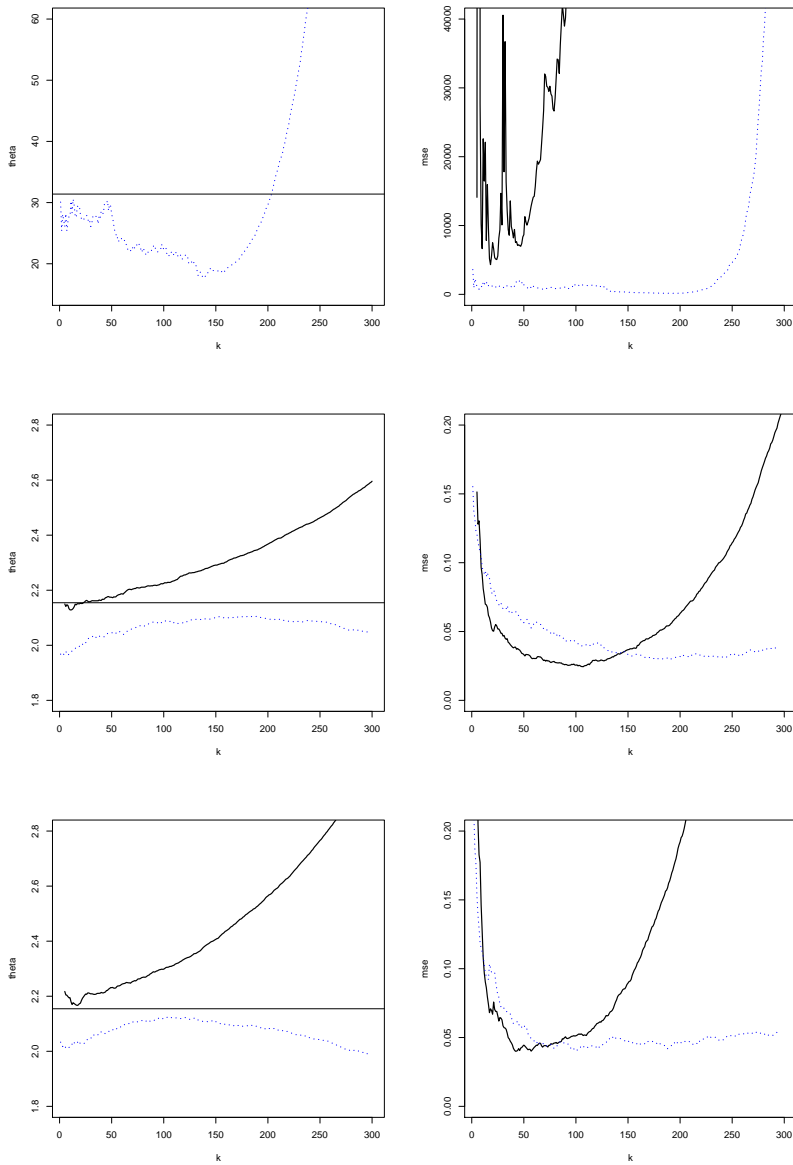
**Fig. F2** Mean (left) and  $MSE$  (right) of  $\hat{\theta}_{1/n,2}^{(H,c)}$  (black solid line) and  $\hat{\theta}_{1/n,2}^{(BC,c)}$  (blue dashed line) as a function of  $k$  for Setup 1 (first row), Setup 2 (second row) and Setup 3 (third row).



**Fig. F3** Mean (left) and  $MSE$  (right) of  $\hat{\Pi}_1^{(c)}(1/n)$  based on  $\hat{\gamma}_k^{(H,c)}$  (black solid line) and  $\hat{\gamma}_k^{(BC,c)}$  (blue dashed line) as a function of  $k$  for Setup 1 (first row), Setup 2 (second row) and Setup 3 (third row).

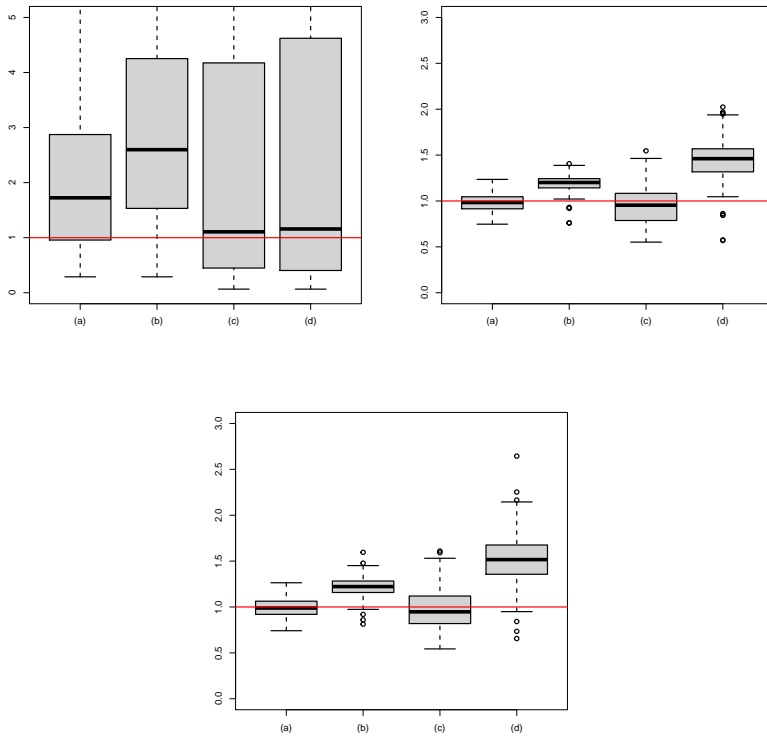


**Fig. F4** Mean (left) and  $MSE$  (right) of  $\hat{V}^{(c)}(1/n)$  based on  $\hat{\gamma}_k^{(H,c)}$  (black solid line) and  $\hat{\gamma}_k^{(BC,c)}$  (blue dashed line) as a function of  $k$  for Setup 1 (first row), Setup 2 (second row) and Setup 3 (third row).



**Fig. F5** Mean (left) and  $MSE$  (right) of  $\hat{\theta}_{1/(1.5n),1}^{(H,c)}$  (black solid line) and  $\hat{\theta}_{1/(1.5n),1}^{(BC,c)}$  (blue dashed line) as a function of  $k$  for Setup 1 (first row), Setup 2 (second row) and Setup 3 (third row).





**Fig. F6** Boxplots of (a)  $\widehat{\theta}_{1/n,1}^{(H,c)}(\widehat{k})/\theta_{1/n,1}$ , (b)  $\widehat{\theta}_{1/n,1}^{(H,c)}(\widetilde{k})/\theta_{1/n,1}$ , (c)  $\widehat{\theta}_{1/n,2}^{(H,c)}(\widehat{k})/\theta_{1/n,2}$  and (d)  $\widehat{\theta}_{1/n,2}^{(H,c)}(\widetilde{k})/\theta_{1/n,2}$  for Setup 1 (top left), Setup 2 (top right) and Setup 3 (bottom).