



HAL
open science

Estimation of the conditional tail moment risk measure under random right censoring

Martin Bladt, Yuri Goegebeur, Armelle Guillou

► **To cite this version:**

Martin Bladt, Yuri Goegebeur, Armelle Guillou. Estimation of the conditional tail moment risk measure under random right censoring. 2025. hal-04879997

HAL Id: hal-04879997

<https://hal.science/hal-04879997v1>

Preprint submitted on 10 Jan 2025

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.

Estimation of the conditional tail moment risk measure under random right 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 and CNRS
7 rue René Descartes, 67084 Strasbourg cedex, France

Abstract

Estimators of the conditional tail moment risk measure based on extreme Kaplan–Meier integral constructions are proposed. The situation when observations are heavy-tailed and subject to right-censoring is considered, which arises often in non-life insurance. Weak convergence is established for both standard and bias-reduced versions of the estimator, and the finite-sample performance is studied through simulations. A real data application to a theft guarantee from a Danish non-life insurer is considered.

Keywords: Conditional tail moment; Pareto-type distribution; Random censorship.

1 Introduction

Claims in non-life insurance often experience long settlement periods, which are further complicated by heavy-tailed claim size distributions and right-

censoring effects. These challenges are particularly present in lines of business with long settlement durations, such as third-party liability insurance. Consequently, the correct estimation of the underlying risks, taking into account the censoring effects while assuming power-law tails has recently been studied in various contexts. For instance, Matthys et al. (2004) considered catastrophic quantile levels, Reynkens et al. (2017) considered parametric distributional fitting, and Bladt et al. (2020) incorporated ultimate information, from individual claims reserving.

In fully observed settings, the study of rare events has been explored extensively through extreme value theory (EVT), which offers a wide range of statistical tools, particularly tailored to answer questions on risk assessment which frequently arise in insurance settings, see, e.g., Chavez-Demoulin & Embrechts (2010) where the importance of EVT for financial risk management is highlighted. In that context, and especially for non-life insurance applications, the most popular risk measure has classically been the Value-at-Risk (VaR) at a level $p \in (0, 1)$, defined as the upper p -th quantile of a loss distribution function F_X associated with a random variable X , i.e., $U_X(1/p)$, where U_X is the tail quantile function of X , that is $U_X(\cdot) := \inf\{x : F_X(x) \geq 1 - \cdot\}$. Despite its versatility and widespread use, it is well known that the VaR suffers from several shortfalls, among them, the fact that it only provides pointwise information, and that it is not a coherent risk measure, as defined by Artzner et al. (1999).

In an attempt to alleviate some of these drawbacks, the conditional tail expectation (CTE), also known as the tail-VaR or expected shortfall in the case of a continuous loss distribution, has been considered extensively by insurers and scholars alike. Its definition in terms of the introduced notation is simply $\mathbb{E}[X|X > U_X(1/p)]$, which is well defined whenever $\mathbb{E}|X| < \infty$. It represents the mean of the loss distribution above a high quantile. Arguably, the CTE takes into account information contained in the entire upper tail of the distribution, making it a more robust tail risk measure. For non-negative random variables, a natural generalization of this measure can then

be formulated by considering further moments $\zeta > 0$ of X , i.e., to consider

$$\theta_{p,\zeta} := \mathbb{E} \left[X^\zeta \mid X > U_X(1/p) \right],$$

which is well defined whenever $\mathbb{E}[X^\zeta] < \infty$, allowing in particular to consider arbitrary moments such as the second moment, and thus the conditional variance of a risk. This risk measure is called the Conditional Tail Moment (CTM, see El Methni et al., 2014), and is of particular interest when data is heavy-tailed.

The goal of this paper is to efficiently estimate $\theta_{p,\zeta}$ not only when the positive random variable X has a heavy-tailed distribution, but when censoring is additionally present. The practical contribution for developing the statistical methodology in this case is tailored for applications in the insurance industry, where claim size distributions are often heavy-tailed, and it is common to encounter censorship for instance when faced with long developments of claims. This is indeed the case in the application we consider, where we analyze ten years of data comprised of claim sizes corresponding to a theft coverage from a Danish non-life insurance company, where settlements take time to be reached. In that case, the correct estimation of the CTM is crucial to understand and assess the risk of any outstanding claim. It is also worth mentioning that our methodology may be used in other application areas, for instance when measuring lifetimes of objects following a power-law distribution. To the best of the authors' knowledge, the only estimator in the literature designed for the same task is the one introduced in Goegebeur et al. (2024), based on a different mathematical construction, which is compared throughout the paper.

The remainder of the paper is organized as follows. In Section 2, we introduce the necessary notation and present recent ideas regarding techniques to handle censored extreme values, which allows us to motivate our new estimator of $\theta_{p,\zeta}$. Then, in Section 3, we establish the main asymptotic properties of our estimator. Subsequently, Section 4 is devoted to a simulation study, where we illustrate the performance of our estimator and compare it with several alternative estimators, including bias-corrected versions. Section 5

presents a real-data application from the non-life insurance sector. All the proofs are postponed to the Appendix.

2 Construction of the estimators adapted to censoring

In this section, we survey the construction of estimators of the CTM adapted to censoring, covering both established methods and new approaches.

2.1 Setting and construction

Throughout this article, X is the variable of interest and its distribution function F_X is assumed to be of the form

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

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 positive measurable function which satisfies

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

The model (1) can be shown to be equivalent to

$$U_X(x) = x^{\gamma_X} \ell_U(x) \quad (2)$$

with $\ell_U(\cdot)$ also a slowly varying function at infinity. Variables having such distribution are said to have regularly varying tails, and are in particular heavy-tailed, in the sense that they do not possess a finite moment generating function.

The variable X is moreover assumed to be censored by another random variable, say Y , assumed to be independent of X , and which also has a heavy-tailed distribution F_Y satisfying

$$F_Y(x) = 1 - x^{-\frac{1}{\gamma_Y}} \ell_Y(x), \quad x > 0, \quad (3)$$

with $\gamma_Y > 0$ and $\ell_Y(\cdot)$ again a slowly varying function at infinity. The censorship framework stipulates that we only observe $Z := \min(X, Y)$ together with an indicator $\delta := \mathbb{1}_{\{X \leq Y\}}$, the latter specifying whether X or Y has been observed. Thus, given a sample $(Z_i, \delta_i)_{\{1 \leq i \leq n\}}$, of independent copies of (Z, δ) , the aim is now to perform inference on $\theta_{p, \zeta}$ when p is very small, i.e., smaller than $1/n$. As mentioned in the introduction, this task has also been considered in Goegebeur et al. (2024). Thus, our objective is to propose an alternative estimator which can outperform the former one in practical situations.

To start with our construction, we notice that using Theorem 1.5.11(ii) in Bingham et al. (1987), for any $p \downarrow 0$ and for $\gamma_X < 1/\zeta$, we have

$$\begin{aligned}
\theta_{p, \zeta} &= -\frac{1}{\overline{F}_X(U_X(1/p))} \int_{U_X(1/p)}^{\infty} x^\zeta d\overline{F}_X(x) \\
&= \left[U_X \left(\frac{1}{p} \right) \right]^\zeta + \frac{\zeta}{\overline{F}_X(U_X(1/p))} \int_{U_X(1/p)}^{\infty} \overline{F}_X(x) x^{\zeta-1} dx \\
&\sim \left[U_X \left(\frac{1}{p} \right) \right]^\zeta \left\{ 1 + \frac{\gamma_X \zeta}{1 - \gamma_X \zeta} \right\} \\
&= \frac{1}{1 - \gamma_X \zeta} \left[U_X \left(\frac{1}{p} \right) \right]^\zeta. \tag{4}
\end{aligned}$$

Consequently, a natural idea is then to estimate $\theta_{p, \zeta}$ based on this asymptotic consideration by

$$\widehat{\theta}_{p, \zeta}^{(c, 1)} := \frac{1}{1 - \widehat{\gamma}_{X, k}^{(c, 1)} \zeta} \left[\widehat{U}_X^{(c, 1)} \left(\frac{1}{p} \right) \right]^\zeta, \tag{5}$$

where k is an intermediate sequence, i.e., a sequence such that $k \rightarrow \infty$ and $\frac{k}{n} \rightarrow 0$, and $\widehat{\gamma}_{X, k}^{(c, 1)}$, resp. $\widehat{U}_X^{(c, 1)}(1/p)$, an estimator of γ_X , resp. for $U_X(1/p)$, adapted to the censorship context. This construction principle was considered in Goegebeur et al. (2024).

An alternative idea to estimate $\theta_{p, \zeta}$ is to use a Weissman-type construction (see Weissman, 1978), which consists in a two-step procedure relying on the following approximation coming from (4)

$$\theta_{p, \zeta} \sim \left[\frac{U_X(\frac{1}{p})}{U_X(\frac{1}{k})} \right]^\zeta \theta_{\frac{k}{n}, \zeta},$$

from which we deduce that

$$\theta_{p,\zeta} \sim \left(\frac{k}{np}\right)^{\zeta \gamma_X} \theta_{\frac{k}{n},\zeta},$$

by Model (2). This leads to a new estimator of $\theta_{p,\zeta}$ defined as

$$\hat{\theta}_{p,\zeta}^{(c,2)} := \left(\frac{k}{np}\right)^{\zeta \hat{\gamma}_{X,k}^{(c,2)}} \tilde{\theta}_{\frac{k}{n},\zeta}^{(c,2)}, \quad (6)$$

where $\hat{\gamma}_{X,k}^{(c,2)}$, resp. $\tilde{\theta}_{k/n,\zeta}^{(c,2)}$, is a suitable estimator of γ_X , resp. for $\theta_{k/n,\zeta}$, adapted to the censorship context. In the case of no censoring, this Weissman-type construction was considered in Goegebeur et al. (2022). The idea is that stabilization can be achieved by the two-step procedure, relying on extrapolation only when considering the far tail. It is also worth noting that the construction principle of (6) does not have the drawback of dividing with a quantity that can possibly become zero, as is the case in (5). Indeed, estimators of extreme value indices in insurance tend to have positive bias when they are based on observations far from the tail, meaning that they estimate γ_X with a value larger than its actual value for medium or large values of k , exacerbating the division problem of (5).

Next, we discuss the estimators of the tail parameters required in both estimators (5) and (6) for $\theta_{p,\zeta}$, that is the estimators of γ_X , $U_X\left(\frac{1}{p}\right)$ and $\theta_{\frac{k}{n},\zeta}$ adapted to censored data.

2.2 Extreme value index estimator in case of censoring

Extreme value theory, and in particular the estimation of the extreme value index for randomly censored data was originally mentioned in Reiss and Thomas (Section 6.1, 1997), where an estimator of a positive extreme value index was introduced without establishing their asymptotic results. Later, this topic was taken up in detail by Beirlant et al. (2007) and Einmahl et al. (2008) for a general extreme value index (i.e., an index in \mathbb{R} instead of a positive index) and for an extreme quantile. In all of these contributions, the idea of the construction of an estimator of γ_X adapted to the censorship framework has been to consider a classical extreme value index estimator

based on the Z -sample, e.g., the Hill estimator (Hill, 1975), and then to divide it by the proportion of non-censored observations in the k -largest Z 's. This leads to

$$\widehat{\gamma}_{X,k}^{(c,1)} = \frac{\frac{1}{k} \sum_{i=1}^k \log Z_{n-i+1,n} - \log Z_{n-k,n}}{\frac{1}{k} \sum_{i=1}^k \delta_{[n-i+1,n]}} \quad (7)$$

where $Z_{i,n}$, $1 \leq i \leq n$, denote the order statistics associated to the Z -sample, and $\delta_{[1,n]}, \dots, \delta_{[n,n]}$ are the δ 's corresponding to $Z_{1,n}, \dots, Z_{n,n}$, respectively.

The idea which motivates this construction is that, due to the independence between X and Y , the observed random variable Z has a distribution function F_Z which satisfies $F_Z(x) = 1 - (1 - F_X(x))(1 - F_Y(x))$, and thus it is also of the form

$$F_Z(x) = 1 - x^{-\frac{1}{\gamma_Z}} \ell_Z(x), \quad x > 0, \quad (8)$$

with $\gamma_Z = \frac{\gamma_X \gamma_Y}{\gamma_X + \gamma_Y} > 0$ and $\ell_Z(\cdot)$ again a slowly varying function at infinity. Thus, the classical Hill estimator based on the Z -sample, i.e., the numerator of (7), estimates γ_Z and not γ_X which is the parameter of interest. As a consequence, this Hill estimator requires adaptation, which can be done by dividing it by the denominator in (7), which represents the proportion of non-censored observations in the k largest Z 's, and thus which estimates $\frac{\gamma_Y}{\gamma_X + \gamma_Y}$. In that way, we recover an estimator of γ_X .

Although the above construction based on moment considerations is easy to conceptualize, it may be sometimes outperformed by newer estimators. For instance, Bladt & Rodionov (2024) studied the asymptotics and favourable finite sample performance of the class of estimators

$$\int_1^\infty \varphi(x) d\mathbb{F}_{k,n}(x), \quad (9)$$

for suitable choices of φ functions, where $\mathbb{F}_{k,n}$ is the extreme Kaplan–Meier estimator defined as

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

In the special case of $\varphi(x) = \log(x)$ the following alternative estimator of γ_X is obtained

$$\hat{\gamma}_{X,k}^{(c,2)} := \int_1^\infty \log(x) d\mathbb{F}_{k,n}(x), \quad (10)$$

which can actually be shown to be equal, after some calculations, to the estimator defined in equation (7) of Worms & Worms (2014). More precisely, this estimator (10) is a consistent estimator in probability for

$$\int_1^\infty \log(x) dF^\circ(x) = \gamma_X,$$

where $F^\circ(x) := \left(1 - x^{-\frac{1}{\gamma_X}}\right) \mathbb{1}_{\{x \geq 1\}}$, as soon as $\gamma_Y > \gamma_X$ (see Section 3 in Blatt & Rodionov, 2024).

2.3 Extreme quantile estimator in case of censoring

Concerning the estimator $\hat{U}_X^{(c,1)}(1/p)$ required in (5), Goegebeur et al. (2024) proposed a Weissman-type construction similar to the one used in (6), but now based on the approximation

$$\begin{aligned} U_X\left(\frac{1}{p}\right) &\sim U_X\left(\frac{n}{k}\right) \left(\frac{k}{np}\right)^{\gamma_X} \\ &\sim U_X\left(\frac{n}{k}\right) \left(\frac{1 - F_X(U_X(\frac{n}{k}))}{p}\right)^{\gamma_X}. \end{aligned} \quad (11)$$

Denoting the Kaplan–Meier product-limit estimator (Kaplan & Meier, 1958) of F_X by

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

Goegebeur et al. (2024) estimate $U_X(1/p)$ by

$$\hat{U}_X^{(c,1)}\left(\frac{1}{p}\right) = Z_{n-k,n} \left(\frac{1 - \mathbb{F}_n(Z_{n-k,n})}{p}\right)^{\hat{\gamma}_{X,k}^{(c,1)}}, \quad (12)$$

where $\hat{\gamma}_{X,k}^{(c,1)}$ is defined in (7).

2.4 Conditional tail moment estimator in case of censoring

We now finally put all the pieces together. We propose an estimator of the form (6) with $\hat{\gamma}_{X,k}^{(c,2)}$ given in (10) and

$$\tilde{\theta}_{\frac{k}{n},\zeta}^{(c,2)} := \left[\hat{U}_X^{(c,2)} \left(\frac{n}{k} \right) \right]^\zeta \int_1^\infty x^\zeta d\mathbb{F}_{k,n}(x) \quad (13)$$

where

$$\hat{U}_X^{(c,2)} \left(\frac{n}{k} \right) := Z_{n-k,n} \left(\frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n} \right)^{\hat{\gamma}_{X,k}^{(c,2)}}. \quad (14)$$

Heuristically, we can understand this estimator as follows: in case of no-censoring $\theta_{\frac{k}{n},\zeta}$ can be estimated by

$$\frac{1}{k} \sum_{i=1}^k X_{n-i+1,n}^\zeta = [X_{n-k,n}]^\zeta \frac{1}{k} \sum_{i=1}^k \left(\frac{X_{n-i+1,n}}{X_{n-k,n}} \right)^\zeta.$$

In case of censoring, according to Bladt & Rodionov (2024), the analogue of

$$\frac{1}{k} \sum_{i=1}^k \left(\frac{X_{n-i+1,n}}{X_{n-k,n}} \right)^\zeta$$

is

$$\int_1^\infty x^\zeta d\mathbb{F}_{k,n}(x),$$

which again conveniently falls into the framework of (9). Concerning the analogue of $X_{n-k,n}$, the natural idea is to use $Z_{n-k,n}$, which is observed, but it should obviously be modified since it estimates $U_Z(n/k)$ and not $U_X(n/k)$. Our modification consists to use a Weissman-type construction similar to the one used in (11), but this time for p replaced by k/n . This yields to estimate the intermediate quantile of X by (14).

In Goegebeur et al. (2024) it was proposed to estimate $\theta_{p,\zeta}$ by (5) with $\hat{\gamma}_{X,k}^{(c,1)}$ and $\hat{U}_X^{(c,1)}(1/p)$ defined in (7) and (12), respectively.

3 Weak convergence of our estimator

When considering the asymptotics of large observations, and in particular when proving consistency of an estimator of a tail parameter, one usually

only needs to assume a first-order condition, i.e., Model (1) suffices. On the contrary, when dealing with the weak convergence of the estimator, one often requires to strengthen this convergence by specifying further the form of the distribution function of X and Y . Since we are interested in this type of asymptotic result, we assume in the sequel that both X and Y satisfy a second-order Pareto-type model given in Assumption (D) below, where RV_ψ denotes 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 $t \rightarrow \infty$, $\forall x > 0$.

Assumption (D) *The survival functions of X and Y satisfy*

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

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

Distribution functions satisfying Assumption (D) also satisfy the so-called second-order condition in univariate extreme value statistics (see, e.g., Theorem 2.3.9 in de Haan and Ferreira, 2006). This assumption is however slightly more general than the Hall class of distribution functions (Hall, 1982, Hall and Welsh, 1985) and is satisfied by many commonly used distributions in insurance, such as the Pareto, Log-gamma, Burr, Fréchet, Student's t-distributions. Under Assumption (D), the tail quantile function $U_\bullet(\cdot)$ satisfies

$$U_\bullet(x) = A_\bullet^{\gamma_\bullet} x^{\gamma_\bullet} (1 + a_\bullet(x)) \tag{15}$$

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 Assumption (D) with $A_Z := A_X A_Y$, $\gamma_Z := \frac{\gamma_X \gamma_Y}{\gamma_X + \gamma_Y}$ and $|\delta_Z(\cdot)| \in RV_{-\beta_Z}$, where $\beta_Z := \min(\beta_X, \beta_Y)$, and hence U_Z satisfies (15).

Throughout the paper, we assume that both $\hat{\gamma}_{X,k}^{(c,i)}$, $i = 1, 2$, defined in (7) and (10) respectively, satisfy the weak convergence

$$\sqrt{k} \left(\hat{\gamma}_{X,k}^{(c,i)} - \gamma_X \right) \xrightarrow{d} \Gamma^{(i)}. \tag{16}$$

This assumption is somewhat immaterial, since under suitable lax assumptions, they are satisfied. The reason for writing these convergences in such a generic fashion is to highlight the fact that *any* estimator satisfying such weak convergence is also a suitable plug-in estimator in our construction for the estimator of the CTM. Such fact becomes clear when examining the proofs in the Appendix. We come back to this point when considering bias-reduced estimators of the extreme value index.

The weak convergence of $\hat{\theta}_{p,\zeta}^{(c,1)}$ defined in (5), has already been established and is repeated below for reference and comparison.

Theorem 3.1 (Theorem 2 in Goegebeur et al., 2024.) Assume (D) with F_X continuous and that the convergence (16) holds for $\hat{\gamma}_{X,k}^{(c,1)}$. 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 a sequence p such that

$$\frac{\log\left(\frac{k}{np^{(\gamma_X + \gamma_Y)/\gamma_Y}}\right)}{\sqrt{k}} \rightarrow 0 \quad \text{and} \quad \frac{k}{np} \rightarrow \infty,$$

we have for $\zeta < 1/\gamma_X$

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

To derive a similar result for our estimator $\hat{\theta}_{p,\zeta}^{(c,2)}$ defined in (6), we need first to establish the weak convergence of the intermediate estimator $\hat{U}_X^{(c,2)}(n/k)$ in (14). This is done in the following theorem.

Theorem 3.2 Assume (D) with F_X continuous and that the convergence (16) holds for $\hat{\gamma}_{X,k}^{(c,2)}$. Then, for $k, n \rightarrow \infty$ and $\sqrt{k} \delta_X(U_Z(n/k)) \rightarrow \lambda \in \mathbb{R}$ and $\sqrt{k}/\log(n/k) \rightarrow \infty$, we have

$$\frac{\sqrt{k}}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \left(\frac{\hat{U}_X^{(c,2)}(n/k)}{U_X(n/k)} - 1 \right) \xrightarrow{d} \Gamma^{(2)}.$$

Note that the condition $\sqrt{k}/\log(n/k) \rightarrow \infty$ is required in order to ensure that the rate of convergence of the above theorem tends to infinity (see (18) in the proof of Theorem 3.2).

Next, we look at the estimator $\tilde{\theta}_{\frac{k}{n}, \zeta}^{(c,2)}$ defined in (13). Its weak convergence is stated below.

Theorem 3.3 *Assume (D) with F_X continuous and that the convergence (16) holds for $\hat{\gamma}_{X,k}^{(c,2)}$. Then, for $k, n \rightarrow \infty$ and $\sqrt{k} \delta_X(U_Z(n/k)) \rightarrow \lambda \in \mathbb{R}$ and $\sqrt{k}/\log(n/k) \rightarrow \infty$, we have when $\zeta \in \left(0, \frac{1}{2} \left[\frac{1}{\gamma_X} - \frac{1}{\gamma_Y} \right] \right)$*

$$\frac{\sqrt{k}}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \left(\frac{\tilde{\theta}_{\frac{k}{n}, \zeta}^{(c,2)}}{\theta_{\frac{k}{n}, \zeta}} - 1 \right) \xrightarrow{d} \zeta \Gamma^{(2)}.$$

We are now able to state the main result of this paper:

Theorem 3.4 *Assume (D) with F_X continuous and that the convergence (16) holds for $\hat{\gamma}_{X,k}^{(c,2)}$. Then, for $k, n \rightarrow \infty$ and $\sqrt{k} \delta_X(U_Z(n/k)) \rightarrow \lambda \in \mathbb{R}$, we have when $\zeta \in \left(0, \frac{1}{2} \left[\frac{1}{\gamma_X} - \frac{1}{\gamma_Y} \right] \right)$*

$$\frac{\sqrt{k}}{\log \frac{k}{np}} \left(\frac{\hat{\theta}_{p, \zeta}^{(c,2)}}{\theta_{p, \zeta}} - 1 \right) \xrightarrow{d} \zeta \left(1 + \frac{\gamma_X}{\gamma_X + \gamma_Y} C \right) \Gamma^{(2)}$$

if $\log(n/k)/\log(k/(np)) \rightarrow C \in [0, \infty)$ and $\sqrt{k}/\log(k/(np)) \rightarrow \infty$, and

$$\frac{\sqrt{k}}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \left(\frac{\hat{\theta}_{p, \zeta}^{(c,2)}}{\theta_{p, \zeta}} - 1 \right) \xrightarrow{d} \zeta \Gamma^{(2)}$$

if $\log(n/k)/\log(k/(np)) \rightarrow \infty$, $\sqrt{k}/\log(n/k) \rightarrow \infty$ and $k/(np) \rightarrow \infty$.

A comparison between Theorems 3.1 and 3.4 leads to the following remarks:

- The rates of convergence of $\hat{\theta}_{p, \zeta}^{(c,1)}$ and $\hat{\theta}_{p, \zeta}^{(c,2)}$, properly normalized, are the same, as well as the limiting distributions (the only difference being $\Gamma^{(i)}$ for $\hat{\theta}_{p, \zeta}^{(c,i)}$, $i = 1, 2$, in the limit), since, according to (18)

$$\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{p} = \frac{\gamma_X}{\gamma_X + \gamma_Y} \log \left(\frac{n}{k} \right) + \log \left(\frac{k}{np} \right) + o_{\mathbb{P}} \left(\log \left(\frac{n}{k} \right) \right)$$

and

$$\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n} = \frac{\gamma_X}{\gamma_X + \gamma_Y} \log \left(\frac{n}{k} \right) + o_{\mathbb{P}} \left(\log \left(\frac{n}{k} \right) \right).$$

- The weak convergence of $\hat{\theta}_{p,\zeta}^{(c,2)}$ is only valid in case $\zeta \in \left(0, \frac{1}{2} \left[\frac{1}{\gamma_X} - \frac{1}{\gamma_Y} \right] \right)$, while that for $\hat{\theta}_{p,\zeta}^{(c,1)}$ requires only $\zeta < 1/\gamma_X$.
- Some of the conditions required in order to have (16) for $\hat{\gamma}_{X,k}^{(c,1)}$, more precisely Conditions (14) and (15) in Einmahl et al. (2008) about $p(z) := \mathbb{P}(\delta = 1|Z = z)$, are not standard conditions and difficult to verify in practice. On the contrary, those required for $\hat{\gamma}_{X,k}^{(c,2)}$ are rather standard (see Appendix E in Bladt & Rodionov, 2024), except the condition $\gamma_Y > \gamma_X$. The latter condition may be practically overcome as explained in Bladt & Rodionov (2024), through the estimation of γ_Z and γ_Y , and retrieving from these two quantities an estimator for γ_X , although weak convergence has not been established in that case.
- It is clear from Theorems 3.1 and 3.4 that the estimators for $\theta_{p,\zeta}$ inherit the behavior of those for γ_X . As it is well known in EVT that the estimation of tail parameters can lead to bias problems, this is the case for the estimators of γ_X and consequently also for those of $\theta_{p,\zeta}$. Bias-reduction has been briefly discussed in Goegebeur et al. (2024), where an estimator introduced by Beirlant et al. (2016) was used instead of $\hat{\gamma}_{X,k}^{(c,1)}$ in (5) and in (12) in order to obtain an asymptotically unbiased estimator of $\theta_{p,\zeta}$. The resulting estimator has a larger asymptotic variance than the initial estimator. However, the estimator of Beirlant et al. (2016) depends on an unknown parameter, which must then be replaced by a canonical value, and thus the bias-correction is generally impractical and lost from a theoretical point of view. Recently, Bladt et al. (2024) introduced a new bias-corrected estimator of γ based on $\hat{\gamma}_{X,k}^{(c,2)}$. This estimator has also the advantage to be a minimum-variance reduced-bias estimator as introduced in Caieiro et al. (2005). As such, it has the same variance as the initial (non-bias-reduced) estimator. This estimator satisfies (16) and thus the above theorem also applies when considering these novel bias-reduced estimators. Both bias-reduced versions of the CTM estimators are also

considered in the next section.

4 A simulation study

In this section, we illustrate the finite sample behavior of our estimator by means of a simulation experiment. We consider the flexible class of $\text{Burr}(\varphi_{\bullet}, \tau_{\bullet})$ distributions, whose distribution function is defined as

$$F_{\bullet}(x) = 1 - (1 + x^{\tau_{\bullet}})^{-\varphi_{\bullet}}, x > 0,$$

for $(\varphi_{\bullet}, \tau_{\bullet}) \in (0, \infty)^2$, where \bullet is either X or Y . This distribution satisfies Assumption (\mathcal{D}) with $\gamma_{\bullet} = 1/(\varphi_{\bullet}\tau_{\bullet})$ and $\beta_{\bullet} = \tau_{\bullet}$. The set of parameters used is $(\varphi_X, \tau_X) = (1, 5)$ which yields $\gamma_X = 0.2$ and $\beta_X = 5$, whereas $(\varphi_Y, \tau_Y) \in \{(1/3.8, 1), (1/1.8, 1), (1/0.8, 1)\}$. These choices yield an asymptotic proportion of censoring $\frac{\gamma_X}{\gamma_X + \gamma_Y}$ of 5%, 10%, and 20%, respectively.

For each distribution, we simulate $N = 1000$ samples of size $n = 500$ and plot different estimators of $\theta_{p,\zeta}$ with $p \in \{\frac{1}{n}, \frac{1}{1.5n}\}$ and $\zeta \in \{1, 2\}$. Due to the restriction $\zeta \in \left(0, \frac{1}{2} \left[\frac{1}{\gamma_X} - \frac{1}{\gamma_Y} \right] \right)$ in Theorem 3.4, the case where $\left(\frac{\gamma_X}{\gamma_X + \gamma_Y}, \zeta\right) = (20\%, 2)$ is not theoretically substantiated for our estimator but kept in order to highlight the robustness of our approach in a case where our theorem does not apply. More precisely, we compare the performance of our estimator $\hat{\theta}_{p,\zeta}^{(c,2)}$ with that of the alternative estimator $\hat{\theta}_{p,\zeta}^{(c,1)}$ introduced in Goegebeur et al. (2024). As explained under Theorem 3.4, these two CTM estimators, correctly normalized, have the same limiting distribution as the extreme value index estimator used in their expressions, and thus may exhibit significant bias effects as is well known in the EVT literature. To illustrate this point, we also compare two bias-corrected versions of these CTM estimators obtained as follows:

- $\hat{\gamma}_{X,k}^{(c,1)}$ is replaced in (5) and (12) by a bias-corrected estimator of γ_X introduced in Beirlant et al. (2016). This yields the estimator $\hat{\theta}_{p,\zeta}^{(cb,1)}$.
- $\hat{\gamma}_{X,k}^{(c,2)}$ is replaced in (6) and (14) by a recent bias-corrected estimator of γ_X proposed in Bladt et al. (2024). The resulting estimator is denoted $\hat{\theta}_{p,\zeta}^{(cb,2)}$.

The algorithms used to compute the two bias-corrected extreme value indices estimators are described in detail in Bladt et al. (2024), where these estimators are also compared and shown to perform well in a simulation study.

In Figure 1 we plot the mean (left) and the MSE (right) based on the N replications when $(\varphi_X, \tau_X) = (1, 5)$ and $(\varphi_Y, \tau_Y) = (1/3.8, 1)$, with, from the top to the bottom, a series of values of (p, ζ) , namely $(\frac{1}{n}, 1)$, $(\frac{1}{n}, 2)$, $(\frac{1}{1.5n}, 1)$ and $(\frac{1}{1.5n}, 2)$. The horizontal lines on the left panel correspond to the true value of $\theta_{p,\zeta}$ computed by numeric integration. Figures 2 and 3 are constructed similarly, but when $(\varphi_Y, \tau_Y) = (1/1.8, 1)$ and $(1/0.8, 1)$, respectively. In terms of mean, the estimators $\hat{\theta}_{p,\zeta}^{(c,i)}$, $i = 1, 2$, are similar, as well as their bias-corrected versions $\hat{\theta}_{p,\zeta}^{(cb,i)}$, but always with a slightly better behavior for our proposals, i.e., when $i = 2$. In terms of MSE, $\hat{\theta}_{p,\zeta}^{(c,2)}$ is always better than $\hat{\theta}_{p,\zeta}^{(c,1)}$ with a much larger improvement when ζ is large. The bias-corrected estimator $\hat{\theta}_{p,\zeta}^{(cb,2)}$ is also the best, with, in Figures 1 and 2, long stability as a function of k compared to $\hat{\theta}_{p,\zeta}^{(cb,1)}$ where the stability is only reached for the highest values of k . This is a favorable property for $\hat{\theta}_{p,\zeta}^{(cb,2)}$ as it implies that the selection of k is not too crucial for this estimator. For a given value of ζ , we can observe that decreasing p complicates the estimation, which is expected since in that case we are interested in a conditional expectation above a more extreme quantile. Conversely, for a given p , increasing ζ implies a greater variability. This is again expected since, in that case, we estimate a higher conditional moment. Finally, by comparing the three figures, we can observe that the higher the proportion of censoring, the more difficult the estimation of $\theta_{p,\zeta}$ is. Nonetheless, our proposals $\hat{\theta}_{p,\zeta}^{(b,2)}$ and $\hat{\theta}_{p,\zeta}^{(cb,2)}$ still perform well in Figure 3, where the restriction on ζ imposed in Theorem 3.4 is not satisfied for the $\zeta = 2$ case.

5 Real data analysis

We consider a dataset of $n = 106\,299$ claims settlement observations corresponding to a theft guarantee of a Danish issuer, with claim arrivals between

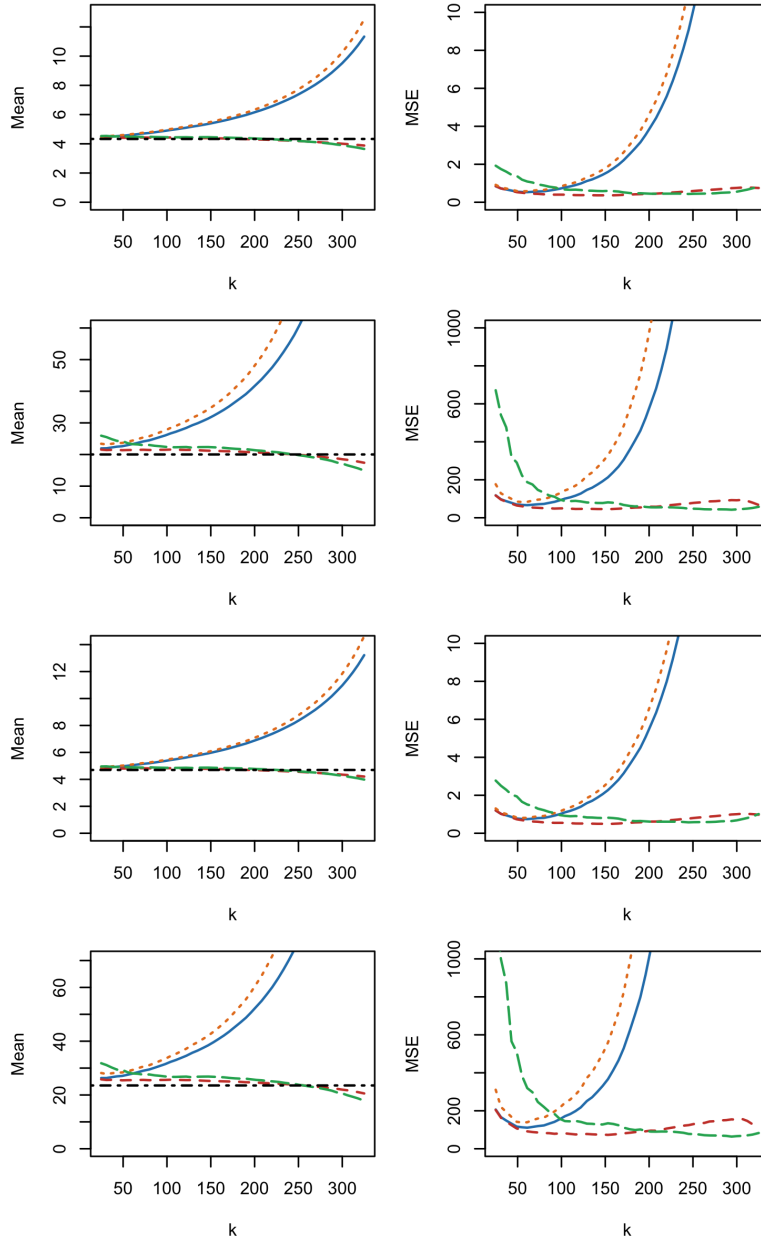


Figure 1: Mean (left) and MSE (right) of $\hat{\theta}_{p,\zeta}^{(c,2)}$ (solid blue line), $\hat{\theta}_{p,\zeta}^{(c,1)}$ (dotted orange line) and the bias-corrected versions $\hat{\theta}_{p,\zeta}^{(cb,2)}$ (dashed red line) and $\hat{\theta}_{p,\zeta}^{(cb,1)}$ (long-dashed green line) based on $N = 1000$ simulations of size $n = 500$, as a function of k . The parameters of the Burr distribution of X are $(1, 5)$ while those for Y correspond to 5% of censoring. From top to bottom: $(p, \zeta) = (\frac{1}{n}, 1), (\frac{1}{n}, 2), (\frac{1}{1.5n}, 1)$ and $(\frac{1}{1.5n}, 2)$.

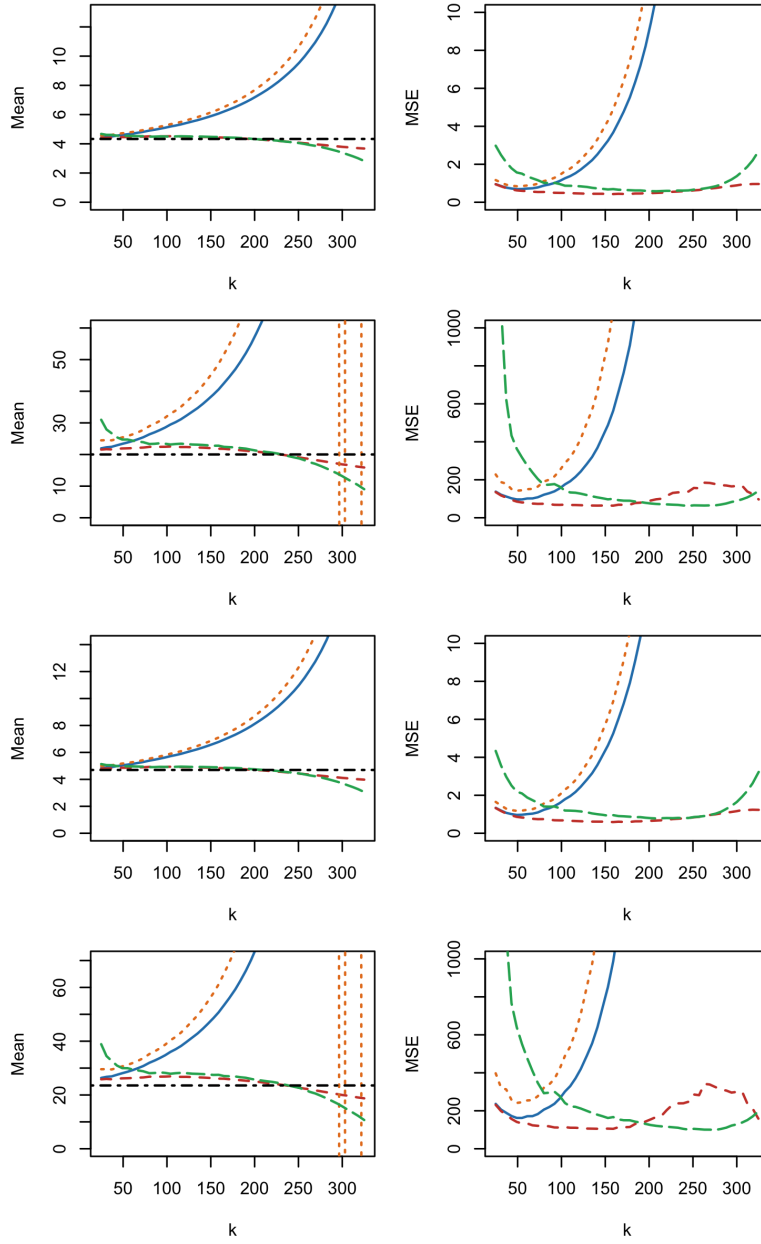


Figure 2: Mean (left) and MSE (right) of $\hat{\theta}_{p,\zeta}^{(c,2)}$ (solid blue line), $\hat{\theta}_{p,\zeta}^{(c,1)}$ (dotted orange line) and the bias-corrected versions $\hat{\theta}_{p,\zeta}^{(cb,2)}$ (dashed red line) and $\hat{\theta}_{p,\zeta}^{(cb,1)}$ (long-dashed green line) based on $N = 1000$ simulations of size $n = 500$, as a function of k . The parameters of the Burr distribution of X are $(1, 5)$ while those for Y correspond to 10% of censoring. From top to bottom: $(p, \zeta) = (\frac{1}{n}, 1), (\frac{1}{n}, 2), (\frac{1}{1.5n}, 1)$ and $(\frac{1}{1.5n}, 2)$.

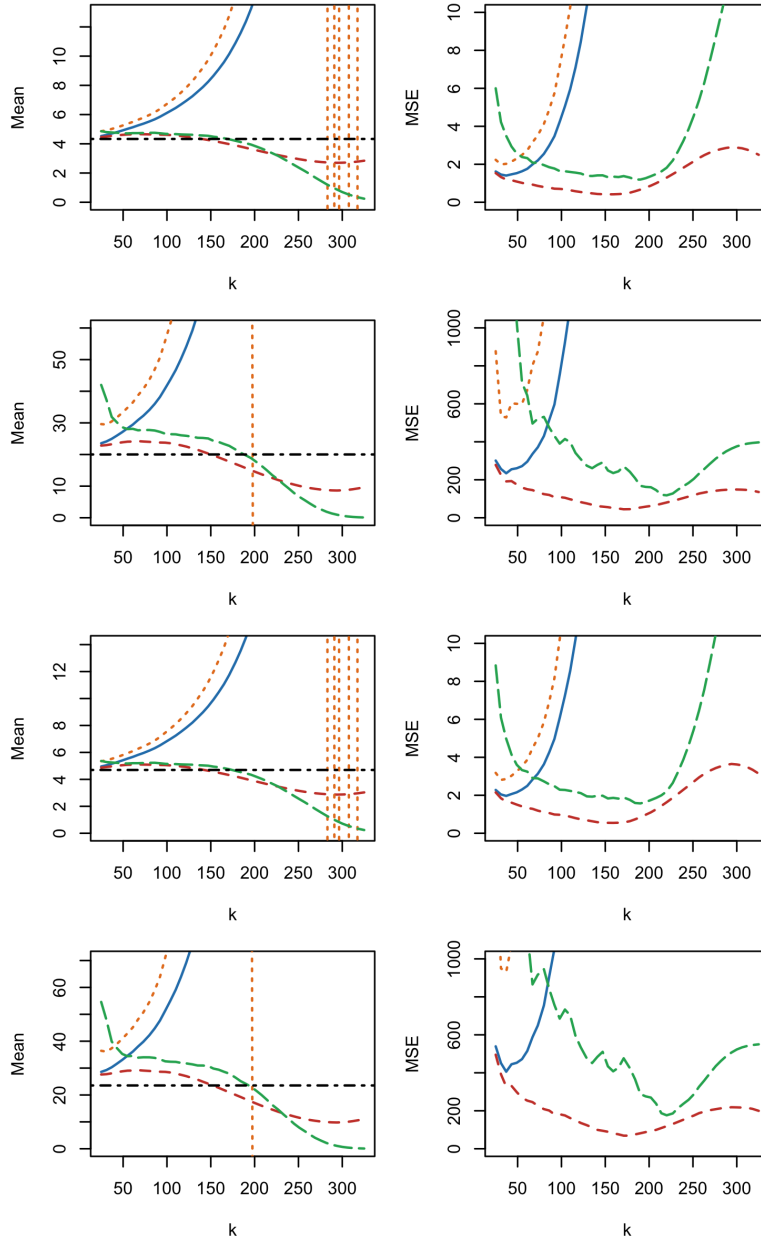


Figure 3: Mean (left) and MSE (right) of $\hat{\theta}_{p,\zeta}^{(c,2)}$ (solid blue line), $\hat{\theta}_{p,\zeta}^{(c,1)}$ (dotted orange line) and the bias-corrected versions $\hat{\theta}_{p,\zeta}^{(cb,2)}$ (dashed red line) and $\hat{\theta}_{p,\zeta}^{(cb,1)}$ (long-dashed green line) based on $N = 1000$ simulations of size $n = 500$, as a function of k . The parameters of the Burr distribution of X are $(1, 5)$ while those for Y correspond to 20% of censoring. From top to bottom: $(p, \zeta) = (\frac{1}{n}, 1), (\frac{1}{n}, 2), (\frac{1}{1.5n}, 1)$ and $(\frac{1}{1.5n}, 2)$.

the 1st of January 2005 and the 31st of December 2014. The theft guarantee of this dataset has not been studied in a statistical context before. Claim sizes develop throughout years, with most settling within just a few months after arrival, but others settlement processes spanning throughout several years. Thus, at the time of data collection, claims are categorized according to *open* or *closed* claims, depending on whether further payments are expected in the future or not, respectively. From a statistical perspective, open claims are right-censored, while close claims are fully observed.

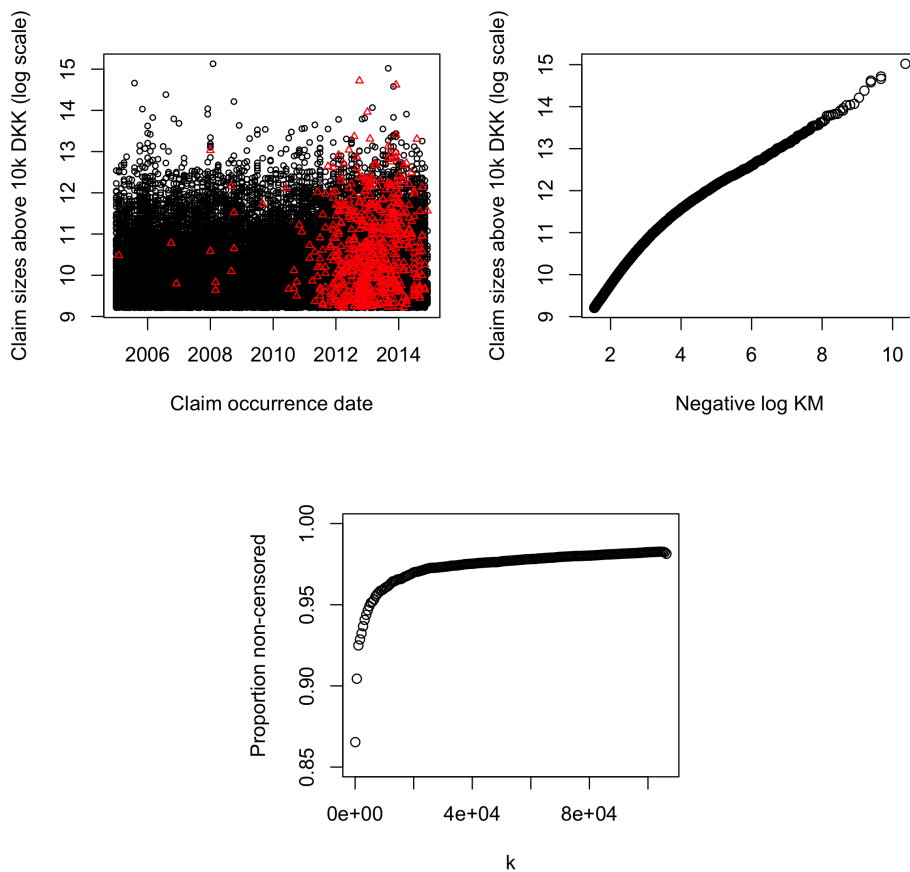


Figure 4: Danish theft insurance. Top left panel: open claim sizes (red triangles) and closed claim sizes (black circles) in order of arrival; top right panel: Kaplan–Meier Pareto quantile plot; bottom panel: empirical proportion of non-censored observations for the k -largest observations.

The dataset is explored in Figure 4. In the top left panel we depict log-claim sizes above 10k DKK (Danish Kroner) in chronological order of arrival, with closed claims represented by black circles and open claims by red triangles. Older claims are less likely to be open, as expected. Also note that claims very close to the evaluation date, 31st of December 2014, tend to not be censored. This is because many claims have arrived but the actual payments are zero in the initial months, which naturally fall below the 10k DKK value.

The validity of the Pareto assumption is checked through the Kaplan–Meier Pareto Quantile plot (see Beirlant et al., 2007) given by

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

and depicted for our dataset in the top right panel of Figure 4. The linear behavior at the largest observations suggests an underlying regularly-varying distribution. We also plot in the bottom panel of Figure 4 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 . We observe stability at the value 98%, and thus our theoretical condition $\gamma_G > \gamma_F$ is satisfied easily.

Next we construct the estimators $\hat{\theta}_{p,\zeta}^{(c,2)}$, $\hat{\theta}_{p,\zeta}^{(c,1)}$, $\hat{\theta}_{p,\zeta}^{(cb,2)}$ and $\hat{\theta}_{p,\zeta}^{(cb,1)}$ as in the simulation study, corresponding to our proposed estimator, the benchmark, and their bias-reduced versions, respectively. The estimates of the corresponding values for γ_F are provided in the top panel of Figure 5, where we observe an agreement between the bias-reduced estimators at a value around 0.4, while the standard (without bias-reduction) estimators have almost no stability. Due to the restriction on ζ in our Theorem 3.4, and the approximated values of the proportion of non-censored observations and of γ_F coming from Figure 4 (bottom) and Figure 5 (top), respectively, our theory can only be applied in case we estimate the CTM with $\zeta = 1$. Thus, for $p = 1/n$ (Figure 5, bottom left) and $p = 1/(1.5n)$ (Figure 5, bottom right) we provide such estimates. We observe a significant discrepancy between the standard and the bias-reduced versions of the estimators, probably due to a value of the second-order parameter $\rho_X = -\beta_X \gamma_X$, estimated here at -0.18 ,

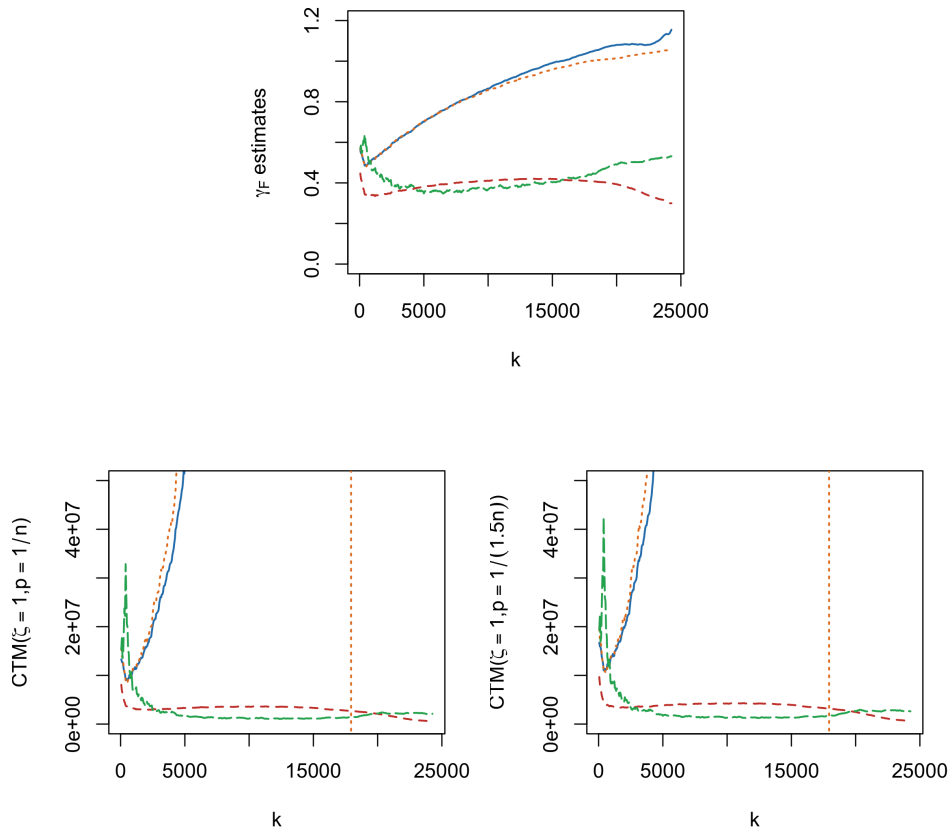


Figure 5: Danish theft insurance. Extreme value index estimators (top) and CTM estimators for $(\zeta, p) = (1, 1/n)$ (bottom left) and for $(\zeta, p) = (1, 1/(1.5n))$ (bottom right). The estimators $\hat{\theta}_{p,\zeta}^{(c,2)}$, $\hat{\theta}_{p,\zeta}^{(c,1)}$, $\hat{\theta}_{p,\zeta}^{(cb,2)}$ and $\hat{\theta}_{p,\zeta}^{(cb,1)}$ are depicted in solid blue, dotted orange, dashed red and long-dashed green respectively, and so are their corresponding γ_F estimates, namely.

too close to 0; We refer to Bladt et al. (2024), where this topic is discussed in details. Finally, note that the differences in values for γ_F estimates are magnified when constructing the CTM risk measures since they appear as a power in (6) and (12). This highlights the delicacy with which the problem must be approached in practice. A combination of expert information and the statistical tools presented in this paper are the best recipe for real-world applications.

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.

Appendix : Proofs

Proof of Theorem 3.2

We use the decomposition

$$\begin{aligned}
\log \frac{\widehat{U}_X^{(c,2)}\left(\frac{n}{k}\right)}{U_X\left(\frac{n}{k}\right)} &= \left(\widehat{\gamma}_{X,k}^{(c,2)} - \gamma_X\right) \log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n} \\
&\quad + \gamma_X \log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{1 - F_X(Z_{n-k,n})} \\
&\quad + \log \left\{ \frac{Z_{n-k,n}}{U_X(n/k)} \left(\frac{1 - F_X(Z_{n-k,n})}{k/n} \right)^{\gamma_X} \right\} \\
&=: \left(\widehat{\gamma}_{X,k}^{(c,2)} - \gamma_X\right) \log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n} + Q_{1,k} + Q_{2,k}.
\end{aligned}$$

Using Csörgő (1996), if F_X is continuous, we have

$$\sup_{x \leq Z_{n-k,n}} \left| \frac{1 - \mathbb{F}_n(x)}{1 - F_X(x)} - 1 \right| = O_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right), \quad (17)$$

from which we deduce that $Q_{1,k} = O_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right)$.

Now, using Assumption (\mathcal{D}) , we have

$$\begin{aligned}
Q_{2,k} &= \gamma_X \log \left(1 + \frac{1}{\gamma_X} \delta_X(Z_{n-k,n}) \right) - \log \left(1 + a_X \left(\frac{n}{k} \right) \right) \\
&= \delta_X(Z_{n-k,n}) (1 + o_{\mathbb{P}}(1)) - \delta_X \left(U_X \left(\frac{n}{k} \right) \right) (1 + o(1)) \\
&= \delta_X \left(U_Z \left(\frac{n}{k} \right) \right) (1 + o_{\mathbb{P}}(1)) - \delta_X \left(U_X \left(\frac{n}{k} \right) \right) (1 + o(1)) \\
&= O \left(\delta_X \left(U_Z \left(\frac{n}{k} \right) \right) \right),
\end{aligned}$$

since $\gamma_Z \leq \gamma_X$. To conclude with the proof of Theorem 3.2, remark that, according to Assumption (\mathcal{D}) and under Model (15), we have using (17)

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

where $\check{Z}_{n-k,n}$ is the $(n-k)$ -largest observation 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

$$\frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n} = A_X^{\frac{\gamma_X}{\gamma_X + \gamma_Y}} A_Y^{-\frac{\gamma_Y}{\gamma_X + \gamma_Y}} \left(\frac{n}{k}\right)^{\frac{\gamma_X}{\gamma_X + \gamma_Y}} (1 + o_{\mathbb{P}}(1)) \xrightarrow{\mathbb{P}} \infty, \quad (18)$$

and as a consequence

$$\frac{\sqrt{k}}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \log \frac{\widehat{U}_X^{(c,2)}(\frac{n}{k})}{U_X(\frac{n}{k})} = \sqrt{k} \left(\widehat{\gamma}_{X,k}^{(c,2)} - \gamma_X \right) + o_{\mathbb{P}}(1) \quad (19)$$

since $\sqrt{k} \delta_X(U_X(n/k)) \rightarrow \lambda \in \mathbb{R}$. This achieves the proof of Theorem 3.2.

Proof of Theorem 3.3

According to Proposition 1 in Goegebeur et al. (2024), we have

$$\frac{\widetilde{\theta}_{\frac{k}{n}, \zeta}^{(c,2)}}{\theta_{\frac{k}{n}, \zeta}} = \left\{ \frac{\widehat{U}_X^{(c,2)}(\frac{n}{k})}{U_X(\frac{n}{k})} \right\}^{\zeta} \frac{\int_1^{\infty} x^{\zeta} d\mathbb{F}_{k,n}(x)}{\frac{1}{1 - \gamma_X \zeta} \{1 + O(\delta_X(U_X(\frac{n}{k})))\}},$$

from which we deduce that

$$\begin{aligned} & \frac{\sqrt{k}}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \left(\frac{\widetilde{\theta}_{\frac{k}{n}, \zeta}^{(c,2)}}{\theta_{\frac{k}{n}, \zeta}} - 1 \right) \\ &= \zeta \frac{\sqrt{k}}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \left(\frac{\widehat{U}_X^{(c,2)}(\frac{n}{k})}{U_X(\frac{n}{k})} - 1 \right) (1 - \gamma_X \zeta) \int_1^{\infty} x^{\zeta} d\mathbb{F}_{k,n}(x) (1 + o_{\mathbb{P}}(1)) \\ &+ \frac{\sqrt{k}}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \left((1 - \gamma_X \zeta) \int_1^{\infty} x^{\zeta} d\mathbb{F}_{k,n}(x) - 1 \right) (1 + o_{\mathbb{P}}(1)) \\ &+ O_{\mathbb{P}} \left(\frac{\sqrt{k} \delta_X(U_X(\frac{n}{k}))}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \right). \end{aligned}$$

The next step consists to prove that

$$\sqrt{k} \left((1 - \gamma_X \zeta) \int_1^{\infty} x^{\zeta} d\mathbb{F}_{k,n}(x) - 1 \right) = O_{\mathbb{P}}(1). \quad (20)$$

To this aim, remark that for $\zeta \in (0, 1/\gamma_X)$, we have

$$\begin{aligned}
& \sqrt{k} \left((1 - \gamma_X \zeta) \int_1^\infty x^\zeta d\mathbb{F}_{k,n}(x) - 1 \right) \\
= & (1 - \gamma_X \zeta) \sqrt{k} \left\{ \int_1^\infty x^\zeta d\mathbb{F}_{k,n}(x) - \int_1^\infty x^\zeta dF^\circ(x) \right\} \\
= & (1 - \gamma_X \zeta) \sqrt{k} \int_1^\infty x^\zeta d \left(\mathbb{F}_{k,n}(x) - F_X^{Z_{n-k,n}}(x) \right) \\
& + (1 - \gamma_X \zeta) \sqrt{k} \int_1^\infty x^\zeta d \left(F_X^{Z_{n-k,n}}(x) - F^\circ(x) \right) \\
= &: Q_{3,k} + Q_{4,k},
\end{aligned}$$

where, for $x \geq 1$,

$$F_X^t(x) := \frac{F_X(tx) - F_X(t)}{1 - F_X(t)}.$$

According to Theorem 3.4 in Bladt & Rodionov, for $\zeta \in \left(0, \frac{1}{2} \left[\frac{1}{\gamma_X} - \frac{1}{\gamma_Y} \right] \right)$

$$\sqrt{k} \int_1^\infty x^\zeta d \left(\mathbb{F}_{k,n}(x) - F_X^{Z_{n-k,n}}(x) \right)$$

is asymptotically normally distributed, and thus $Q_{3,k} = O_{\mathbb{P}}(1)$.

Now, we need to handle $Q_{4,k}$. To this aim, we consider a random variable ξ following a strict unit Pareto distribution and being independent of the Z -sample. Since

$$\mathbb{P} \left(\frac{U_X(s\xi)}{U_X(s)} \leq x \right) = \frac{F_X(U_X(s)x) - F_X(U_X(s))}{1 - F_X(U_X(s))},$$

we deduce that F_X^t is the distribution function of $\frac{U_X(U_X^-(t)\xi)}{U_X(U_X^-(t))}$. Using the fact that $U_X^-(t) = [1 - F_X(t)]^{-1}$, we deduce that

$$Q_{4,k} = (1 - \gamma_X \zeta) \sqrt{k} \left\{ \mathbb{E} \left[\left(\frac{U_X([1 - F_X(Z_{n-k,n})]^{-1}\xi)}{U_X([1 - F_X(Z_{n-k,n})]^{-1})} \right)^\zeta \right] - \mathbb{E}[\xi^{\gamma_X \zeta}] \right\}.$$

Now, for a fixed s , using a Taylor expansion, we have

$$\mathbb{E} \left[\left(\frac{U_X(s\xi)}{U_X(s)} \right)^\zeta \right] - \mathbb{E}[\xi^{\gamma_X \zeta}] = \zeta \mathbb{E} \left[\{u(s)\}^{\zeta-1} \left\{ \frac{U_X(s\xi)}{U_X(s)} - \xi^{\gamma_X} \right\} \right],$$

where $u(s)$ is an intermediate random value between $\frac{U_X(s\xi)}{U_X(s)}$ and ξ^{γ_X} . Since, under Model (15), we have

$$\lim_{s \rightarrow \infty} \frac{\frac{U_X(sx)}{U_X(s)} - x^{\gamma_X}}{-\gamma_X \beta_X \frac{a_X(s)}{1+a_X(s)}} = x^{\gamma_X} \frac{x^{-\gamma_X \beta_X} - 1}{-\gamma_X \beta_X},$$

according to Theorem 2.3.9 in de Haan and Ferreira (2006), for all $\varepsilon, \delta > 0$ there exists $s_0 = s_0(\varepsilon, \delta)$ such that for all $s \geq s_0$ and $x \geq 1$ we have

$$\left| \frac{\frac{U_X(sx)}{U_X(s)} - x^{\gamma_X}}{-\gamma_X \beta_X A_X(s)} - x^{\gamma_X} \frac{x^{-\gamma_X \beta_X} - 1}{-\gamma_X \beta_X} \right| \leq \varepsilon x^{\gamma_X - \gamma_X \beta_X + \delta},$$

with $A_X \sim a_X$. Consequently

$$\begin{aligned} \mathbb{E} \left[\left(\frac{U_X(s\xi)}{U_X(s)} \right)^\zeta \right] - \mathbb{E}[\xi^{\gamma_X \zeta}] &= \zeta A_X(s) \mathbb{E} \left[\{u(s)\}^{\zeta-1} \xi^{\gamma_X} \{\xi^{-\gamma_X \beta_X} - 1\} \right] \\ &\quad + o_{\mathbb{P}} \left(A_X(s) \mathbb{E} \left[\{u(s)\}^{\zeta-1} \xi^{\gamma_X - \gamma_X \beta_X + \delta} \right] \right) \\ &= O_{\mathbb{P}}(A_X(s)) \end{aligned}$$

since combining Potter's inequalities (see, e.g., de Haan & Ferreira, Proposition B.1.9.5, 2006) with the dominated convergence theorem, we have

$$\begin{aligned} \mathbb{E} \left[\{u(s)\}^{\zeta-1} \xi^{\gamma_X} \{\xi^{-\gamma_X \beta_X} - 1\} \right] &\longrightarrow \mathbb{E} \left[\xi^{\gamma_X \zeta} \{\xi^{-\gamma_X \beta_X} - 1\} \right] < \infty \\ \mathbb{E} \left[\{u(s)\}^{\zeta-1} \xi^{\gamma_X - \gamma_X \beta_X + \delta} \right] &\longrightarrow \mathbb{E} \left[\xi^{\gamma_X \zeta - \gamma_X \beta_X + \delta} \right] < \infty. \end{aligned}$$

This yields

$$\begin{aligned} Q_{4,k} &= O_{\mathbb{P}} \left(\sqrt{k} A_X \left([1 - F_X(Z_{n-k,n})]^{-1} \right) \right) \\ &= O_{\mathbb{P}} \left(\sqrt{k} \delta_X \left(U_X \left([1 - F_X(Z_{n-k,n})]^{-1} \right) \right) \right) \\ &= O_{\mathbb{P}} \left(\sqrt{k} \delta_X \left(Z_{n-k,n} \right) \right) \\ &= O_{\mathbb{P}} \left(\sqrt{k} \delta_X \left(U_Z \left(\frac{n}{k} \right) \right) \right) \\ &= O_{\mathbb{P}}(1) \end{aligned}$$

which ends the proof of (20). Finally using (19), we deduce that

$$\frac{\sqrt{k}}{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}} \left(\frac{\tilde{\theta}_{\frac{k}{n}, \zeta}^{(c,2)}}{\theta_{\frac{k}{n}, \zeta}} - 1 \right) = \zeta \sqrt{k} \left(\hat{\gamma}_{X,k}^{(c,2)} - \gamma_X \right) + o_{\mathbb{P}}(1), \quad (21)$$

from which Theorem 3.3 follows.

Proof of Theorem 3.4

We use the decomposition

$$\begin{aligned}
\frac{\widehat{\theta}_{p,\zeta}^{(c,2)}}{\theta_{p,\zeta}} - 1 &= \left\{ \frac{\widetilde{\theta}_{\frac{k}{n},\zeta}^{(c,2)}}{\theta_{\frac{k}{n},\zeta}} - 1 \right\} \frac{\theta_{\frac{k}{n},\zeta}}{\theta_{p,\zeta}} \left(\frac{k}{np} \right)^{\zeta \widehat{\gamma}_{X,k}^{(c,2)}} \\
&\quad + \left\{ \frac{\theta_{\frac{k}{n},\zeta}}{\theta_{p,\zeta}} \left(\frac{k}{np} \right)^{\zeta \gamma_X} - 1 \right\} \left(\frac{k}{np} \right)^{\zeta (\widehat{\gamma}_{X,k}^{(c,2)} - \gamma_X)} \\
&\quad + \left\{ \left(\frac{k}{np} \right)^{\zeta (\widehat{\gamma}_{X,k}^{(c,2)} - \gamma_X)} - 1 \right\} \\
&=: Q_{5,k} \frac{\theta_{\frac{k}{n},\zeta}}{\theta_{p,\zeta}} \left(\frac{k}{np} \right)^{\zeta \widehat{\gamma}_{X,k}^{(c,2)}} + Q_{6,k} \left(\frac{k}{np} \right)^{\zeta (\widehat{\gamma}_{X,k}^{(c,2)} - \gamma_X)} + Q_{7,k}.
\end{aligned}$$

Using Proposition 1 in Goegebeur et al. (2024), we have under Model (15) and since $k/(np) \rightarrow \infty$

$$\begin{aligned}
Q_{6,k} &= \left(\frac{1 + a_X\left(\frac{n}{k}\right)}{1 + a_X\left(\frac{1}{p}\right)} \right)^{\zeta} \frac{1}{1 - \gamma_X \zeta} + O(\delta_X(U_X\left(\frac{n}{k}\right))) - 1 \\
&= O\left(\delta_X\left(U_X\left(\frac{n}{k}\right)\right)\right) + O\left(\delta_X\left(U_X\left(\frac{1}{p}\right)\right)\right) \\
&= O\left(\delta_X\left(U_X\left(\frac{n}{k}\right)\right)\right) \\
&= O\left(\delta_X\left(U_Z\left(\frac{n}{k}\right)\right)\right). \tag{22}
\end{aligned}$$

Also, under the condition $\log(k/(np))/\sqrt{k} \rightarrow 0$, we have

$$Q_{7,k} = \zeta \left(\widehat{\gamma}_{X,k}^{(c,2)} - \gamma_X \right) \log \frac{k}{np} (1 + o_{\mathbb{P}}(1)). \tag{23}$$

Combining (21) with (22) and (23), we deduce that

$$\begin{aligned}
\frac{\widehat{\theta}_{p,\zeta}^{(c,2)}}{\theta_{p,\zeta}} - 1 &= \zeta \left(\widehat{\gamma}_{X,k}^{(c,2)} - \gamma_X \right) \left\{ \log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n} + \log \frac{k}{np} \right\} \\
&\quad + o_{\mathbb{P}} \left(\frac{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}}{\sqrt{k}} \right) + o_{\mathbb{P}} \left(\frac{\log \frac{k}{np}}{\sqrt{k}} \right) + O\left(\delta_X\left(U_Z\left(\frac{n}{k}\right)\right)\right).
\end{aligned}$$

Finally, since according to (18), the condition

$$\frac{\log \frac{1 - \mathbb{F}_n(Z_{n-k,n})}{k/n}}{\log \frac{k}{np}} \xrightarrow{\mathbb{P}} D \in [0, \infty]$$

is equivalent to the condition

$$\frac{\log \frac{n}{k}}{\log \frac{k}{np}} \longrightarrow \frac{\gamma_X + \gamma_Y}{\gamma_X} D \in [0, \infty],$$

since $k/(np) \rightarrow \infty$, Theorem 3.4 follows.

References

- Artzner P., Delbaen F., Eber, J.M. & Heath D. (1999). Coherent measures of risk, *Mathematical Finance*, 9, 203–228.
- Beirlant, J., Bardoutsos, A., de Wet, T. & Gijbels, I. (2016). Bias reduced tail estimation for censored Pareto type distributions, *Statistics & Probability Letters*, 109,78–88.
- Beirlant, J., Guillou, A., Dierckx, G. & Fils-Villetard, A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring, *Extremes*, 10, 151–174.
- Bingham, N.H., Goldie, C.M. & Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press.
- Bladt, M., Albrecher, H. & Beirlant, J. (2020). Combined tail estimation using censored data and expert information, *Scandinavian Actuarial Journal*, 2020(6), 503–525.
- Bladt M., Goegebeur, Y. & Guillou, A. (2024). Asymptotically unbiased estimator of the extreme value index under random censoring, <https://hal.science/hal-04786783>.
- Bladt M. & Rodionov, I. (2024). Censored extreme value estimation, <https://arxiv.org/abs/2312.10499>.
- Caeiro, F., Gomes, M.I. & Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator, *REVSTAT*, 113–136.
- Chavez-Demoulin, V. & Embrechts, P. (2010). Revisiting the edge, ten years on, *Communications in Statistics - Theory & Methods*, 39, 1674–1688.
- Csörgő, S. (1996). Universal Gaussian approximations under random censorship, *The Annals of Statistics*, 24, 2744–2778.
- Einmahl, J.H.J., Fils-Villetard, A. & Guillou, A. (2008). Statistics of extremes under random censoring, *Bernoulli*, 14, 207–227.
- El Methni, J., Gardes, L. & Girard, S. (2014). Non-parametric estimation of extreme risk measures from conditional heavy-tailed distributions, *Scandinavian Journal of Statistics*, 41(4), 988–1012.
- Goegebeur, Y., Guillou, A., Pedersen, T. & Qin, J. (2022). Extreme-

value based estimation of the conditional tail moment with application to reinsurance rating, *Insurance: Mathematics and Economics*, 107, 102–122.

Goegebeur, Y., Guillou, A. & Qin, J. (2024). Conditional tail moment and reinsurance premium estimation under random right censoring, *TEST*, 33, 230–250.

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

Hall, P. (1982). On some simple estimates of an exponent of regular variation, *Journal of the Royal Statistical Society - Serie B*, 44, 37–42.

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

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

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

Matthys, G., Delafosse, E., Guillou, A. & Beirlant, J. (2004). Estimating catastrophic quantile levels for heavy-tailed distributions, *Insurance: Mathematics and Economics*, 34(3), 517–537.

Reiss, R.D. & Thomas, M. (1997). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Basel, Birkhäuser.

Reynkens, T., Verbelen, R., Beirlant, J. & Antonio, K. (2017). Modelling censored losses using splicing: A global fit strategy with mixed Erlang and extreme value distributions, *Insurance: Mathematics and Economics*, 77, 65–77.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *Journal of the American Statistical Association*, 73, 812–815.

Worms, J. & Worms, R. (2014). New estimators of the extreme value index under random right censoring, for heavy-tailed distributions, *Extremes*,

17, 337–358.