



# Estimating extreme quantiles under random truncation

Laurent Gardes, Gilles Stupfler

## ► To cite this version:

Laurent Gardes, Gilles Stupfler. Estimating extreme quantiles under random truncation. *Test*, 2015, 24 (2), pp.207-227. 10.1007/s11749-014-0403-5 . hal-00942134v2

**HAL Id: hal-00942134**

**<https://hal.science/hal-00942134v2>**

Submitted on 16 Sep 2014

**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.

# Estimating extreme quantiles under random truncation

Laurent Gardes<sup>(1,\*)</sup> and Gilles Stupfler<sup>(2)</sup>

<sup>(1)</sup> Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes,  
67084 Strasbourg Cedex, France.

\* Corresponding author, [gardes@unistra.fr](mailto:gardes@unistra.fr)

<sup>(2)</sup> Aix Marseille Université, CNRS, EHESS, Centrale Marseille, GREQAM UMR  
7316, 13002 Marseille, France

## Abstract

The goal of this paper is to provide estimators of the tail index and extreme quantiles of a heavy-tailed random variable when it is right-truncated. The weak consistency and asymptotic normality of the estimators are established. The finite sample performance of our estimators is illustrated on a simulation study and we showcase our estimators on a real set of failure data. **keywords:** Asymptotic normality, consistency, extreme quantile, heavy-tailed distribution, tail index. **AMS Subject Classifications:** 62G05, 62G20, 62G30, 62G32.

## 1 Introduction

Studying extreme events is relevant in numerous fields of statistical applications. One can think about hydrology, where one may want to estimate the maximum level reached by seawater along a coast over a given period, or to study extreme rainfall at a given location; in actuarial science, a pivotal problem for an insurance firm is to estimate the probability that a claim so large that it represents a threat to its solvency is filed. In this type of problem, the focus is not in the estimation of “central” parameters of the random variable of interest, such as its mean or median, but rather in the understanding of its behavior in its right tail.

A particular relevant case is when the random variable of interest,  $Y$ , is heavy-tailed, namely, when its survival function  $\bar{F}$  can be written  $\bar{F}(y) = y^{-1/\gamma} L(y)$  for all  $y > 0$ ; here,  $\gamma > 0$  shall be referred to as the tail index and  $L$  is a slowly varying function at infinity, meaning that it satisfies  $L(\lambda y)/L(y) \rightarrow 1$  as  $y \rightarrow \infty$  for all  $\lambda > 0$ . In this case,  $\gamma$  clearly drives the tail behavior of  $\bar{F}$  and its knowledge is necessary if, for instance, we are interested in the estimation of extreme quantiles of  $Y$ . The estimation of the tail index is thus one of the central topics in extreme value theory, which is why this problem has been extensively studied

in the literature. Recent overviews on univariate tail index estimation can be found in the monographs of Beirlant et al. (2004) and de Haan and Ferreira (2006).

A further challenge arises when facing incomplete data. An example of such a situation is the estimation of (extreme) survival times based on a follow-up study of patients suffering from a given illness. If at the time the data are collected a patient is still alive, then his/her survival time is not available to the researcher, although it is known that the patient survived until the end of the study. This case is the archetypal example of right-censoring. Estimating the tail index in this situation is much more difficult than when having complete data, since information about the right tail of the variable of interest is missing. In this setting, the estimation of the tail index and extreme quantiles has been considered by Beirlant and Guillou (2001), Beirlant et al. (2007), Beirlant et al. (2010), Einmahl et al. (2008), Gomes and Neves (2011) and Worms and Worms (2014).

In this paper, we consider the case when the data are right-truncated. In this framework, one observes the variable of interest if and only if it is less than or equal to a truncation variable  $T$ . This situation is different from right-censoring since nothing is known about  $Y$  in the case  $Y > T$ , which adds a further difficulty to the analysis of the right tail of  $Y$ . Truncated data may be collected in various cases, for instance when estimating incubation times for a given disease, see Kalbfleisch and Lawless (1989, 1991) and Lagakos et al. (1988); when studying the luminosity of astronomical objects such as quasars, see Jackson (1974) and Lynden-Bell (1971); when accounting for reporting lags in insurance data, also referred to as the incurred but not yet reported problem, see Herbst (1999), Kaminsky (1987) and Lawless (1994); or when considering failure or warranty data, see Hu and Lawless (1996a, 1996b) and the monographs by Meeker and Escobar (1998) and Lawless (2002). To the best of our knowledge, the estimation of the tail index and extreme quantiles in this context is, up to now, still an open question.

The outline of this paper is as follows. In Section 2, we give a precise definition of our model and define our estimators of the tail index and of the extreme quantiles of a truncated random variable. Some asymptotic properties of our estimators are stated in Section 3. The finite sample performance of the extreme quantile estimator is studied in Section 4. A real set of brake failure data is analyzed in Section 5. We offer some concluding remarks in Section 6. Proofs of the main results are deferred to Section 7, while the preliminary results are deferred to the Appendix. Proofs of the preliminary results can be found in a supplementary material.

## 2 Framework

Let  $(Y_1, T_1), \dots, (Y_n, T_n)$  be  $n$  independent copies of a random pair  $(Y, T) \in [y_0, \infty) \times [t_0, \infty)$ , where  $Y$  and  $T$  are independent and  $y_0, t_0 \geq 0$  are the left endpoints of  $Y$  and  $T$ . The right endpoints of  $Y$  and  $T$  are supposed to be

infinite. The joint cumulative distribution function (cdf) of the random pair  $(Y, T)$  is then given for all  $(y, t) \in \mathbb{R}^2$  by  $H(y, t) := \mathbb{P}(Y \leq y, T \leq t) = F(y)G(t)$ , where  $F$  and  $G$  are the cdfs of  $Y$  and  $T$ . The focus of this paper is on extreme quantiles of  $Y$  and, as a first step, on the estimation of the cdf  $F$ . Of course, because we only record the  $Y_i$  and  $T_i$  such that  $Y_i \leq T_i$ , the classical nonparametric estimator of  $F$  cannot be used. However, the conditional cdfs of  $Y$  and  $T$  given  $Y \leq T$ , respectively denoted by  $F^*$  and  $G^*$ , may be estimated in a nonparametric way. Let  $N$  be the total (random) number of observed pairs  $(Y_i, T_i)$  such that  $Y_i \leq T_i$  and notice that  $N$  is a binomial random variable with parameters  $n$  and  $p := \mathbb{P}(Y \leq T)$ . Such pairs shall be denoted in the sequel as  $(Y_i^*, T_i^*)$ ,  $1 \leq i \leq N$ . It can be shown (see Lemma 1) that the conditional distribution of  $\{(Y_i^*, T_i^*), i = 1, \dots, N\}$  given  $N = m$  is equal to the distribution of  $m$  independent copies of a random vector  $(Y^*, T^*)$  with cdf  $H^*$  given by  $H^*(y, t) = \mathbb{P}(Y \leq y, T \leq t | Y \leq T)$ . The standard estimators of the conditional cdfs of  $Y$  and  $T$  are then

$$\hat{F}_N^*(y) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{\{Y_i^* \leq y\}} \text{ and } \hat{G}_N^*(t) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{\{T_i^* \leq t\}}.$$

Note now that, in order to estimate  $F$ , it is sufficient to estimate the function  $\Lambda^F := -\log F$ . The following result, whose proof can be found in Woodroffe (1985, p.166), shows that this quantity is in fact linked to  $F^*$  and  $G^*$ :

**Proposition 1** *Let  $C^* := F^* - G^*$ . Then  $C^*(y) = p^{-1}F(y)(1 - G(y)) > 0$  for all  $y > y_0$ , and*

$$\Lambda^F(y) = \int_y^\infty \frac{dF(z)}{F(z)} = \int_y^\infty \frac{dF^*(z)}{C^*(z)}.$$

This result can be used to build an estimator of the function  $\Lambda^F$  and consequently of the cdf  $F$ : if  $y > y_0$ , we may estimate  $\Lambda^F(y)$  by

$$\hat{\Lambda}_N^F(y) = \frac{1}{N} \sum_{i=1}^N \frac{\mathbb{I}_{\{Y_i^* > y\}}}{\hat{C}_N^*(Y_i^*)}$$

when  $N > 0$  and 0 otherwise. The survival function  $\bar{F} := 1 - F$  and its associated quantile function  $\alpha \mapsto q(\alpha) := \inf\{y \geq y_0 | \bar{F}(y) \leq \alpha\}$ , which is the right-continuous inverse of  $\bar{F}$ , are then estimated by  $\hat{F}_N(y) = \exp(-\hat{\Lambda}_N^F(y))$  and  $\hat{q}_N(\alpha) = \inf\{y \geq y_0 | \hat{F}_N(y) \leq \alpha\}$ , where we let  $\hat{\bar{F}}_N = 1 - \hat{F}_N$ . The first aim of this paper is to study the asymptotic behavior of the estimator  $\hat{q}_N(\alpha_n)$  where  $\alpha_n \rightarrow 0$  as  $n \rightarrow \infty$ . We shall tackle this problem in a framework of regular variation: we write that a function  $\Psi \in \mathcal{RV}_1(a)$ ,  $a \in \mathbb{R}$ , if  $\Psi$  is nonnegative on  $(0, \infty)$  and for all  $\lambda > 0$ , we have  $\Psi(\lambda y)/\Psi(y) \rightarrow \lambda^a$  as  $y \rightarrow \infty$ . We thus consider the following model:

**(M)** We have  $\bar{F} \in \mathcal{RV}_1(-1/\gamma_F)$  and  $\bar{G} \in \mathcal{RV}_1(-1/\gamma_G)$ , with  $0 < \gamma_F \leq \gamma_G$ .

Model **(M)** is a standard extreme-value model adapted to right-truncated data; see also Beirlant et al. (2007) and Einmahl et al. (2008) for closely related models when there is right-censoring. The cdfs of  $Y$  and  $T$  are thus heavy-tailed with respective tail indices  $\gamma_F$  and  $\gamma_G$ . The condition  $\gamma_F \leq \gamma_G$  ensures that we have at our disposal enough observations pertaining to the right tail of  $Y$ . In this context, the quantile estimator  $\hat{q}_N(\alpha_n)$  is consistent if  $\alpha_n \rightarrow 0$  slowly enough, see Theorem 2. In order to remove the restriction on the rate of convergence of  $\alpha_n$ , we note that under model **(M)**, the quantile function  $q$  is regularly varying at 0 (see Corollary 1.2.10 p.23 in de Haan and Ferreira 2006), so that if  $\beta_n < \alpha_n$  are two positive sequences tending to 0 such that  $\beta_n/\alpha_n \rightarrow 0$  then  $q(\beta_n) \approx q(\alpha_n) (\alpha_n/\beta_n)^{\gamma_F}$  when  $n$  is large. In order to derive an estimator of an extreme quantile  $q(\beta_n)$  from that, we first need to build an estimator of  $\gamma_F$ . With this aim in mind, we remark that  $\bar{F}^*$  and  $\bar{G}^*$  are heavy-tailed with respective tail indices  $\gamma_{F^*} := \gamma_F \gamma_G / (\gamma_F + \gamma_G)$  and  $\gamma_G$  (see Lemma 3) and we introduce the Hill-type estimators (see Hill 1975)

$$\hat{\gamma}_{N,F^*}(k_N) = \frac{1}{k_N} \sum_{i=1}^{k_N} \log \frac{Y_{N-i+1,N}^*}{Y_{N-k_N,N}^*} \text{ and } \hat{\gamma}_{N,G}(k'_N) = \frac{1}{k'_N} \sum_{i=1}^{k'_N} \log \frac{T_{N-i+1,N}^*}{T_{N-k'_N,N}^*}.$$

Here we let, given  $N = m$ ,  $k_N = k_m$  and  $k'_N = k'_m$ , where  $k_m$  and  $k'_m$  are integers which belong to  $\{1, \dots, m-1\}$ , and  $Y_{1,N}^* \leq \dots \leq Y_{N,N}^*$ ,  $T_{1,N}^* \leq \dots \leq T_{N,N}^*$  are the order statistics deduced from the samples  $(Y_i^*)_{1 \leq i \leq N}$ ,  $(T_i^*)_{1 \leq i \leq N}$ . It is fairly easy to prove (see Lemma 9) that  $\hat{\gamma}_{N,F^*}(k_N)$  and  $\hat{\gamma}_{N,G}(k'_N)$  are consistent estimators of  $\gamma_{F^*}$  and  $\gamma_G$  under mild conditions. This leads us to introduce the class of estimators

$$\hat{\gamma}_{N,F}(k_N, k'_N) = \frac{\hat{\gamma}_{N,F^*}(k_N) \hat{\gamma}_{N,G}(k'_N)}{\hat{\gamma}_{N,G}(k'_N) - \hat{\gamma}_{N,F^*}(k_N)}. \quad (1)$$

Under some conditions on  $(k_m)$  and  $(k'_m)$ , the quantity  $\hat{\gamma}_{N,F}(k_N, k'_N)$  is then a consistent estimator of  $\gamma_F$ , see Theorem 3. This motivates the following Weissman-type estimator (see Weissman 1978) for a quantile having arbitrary order  $\beta_n \rightarrow 0$ :

$$\hat{q}_N^W(\beta_n | \alpha_n, k_N, k'_N) = \hat{q}_N(\alpha_n) (\alpha_n/\beta_n)^{\hat{\gamma}_{N,F}(k_N, k'_N)} \quad (2)$$

where  $\alpha_n \rightarrow 0$  converges slowly enough. The asymptotic properties of this estimator are discussed in Theorem 4.

### 3 Main results

In this section, we examine the asymptotic properties of our estimators. In order to establish the asymptotic normality of  $\hat{\bar{F}}_N(y_n)$ , we introduce the following additional condition:

$$\int_{y_0}^{\infty} \frac{dF(z)}{G(z)} < \infty. \quad (3)$$

This assumption is classical in the study of the estimator of the cdf of a truncated random variable, see for instance Stute and Wang (2008) and Woodroffe (1985) for related hypotheses when there is left-truncation. Note that under model **(M)**, it is a consequence of Lemma 2 with  $\varphi = 1/\overline{G}$  and  $\psi = \overline{F}$  that (3) automatically holds if  $\gamma_F < \gamma_G$ . Besides, it is easy to check that (3) fails to hold if  $\gamma_F > \gamma_G$ .

**Theorem 1** *Let  $y_n \rightarrow \infty$ . Assume that **(M)** and (3) hold, and that  $nv^2(y_n)$  converges to infinity where*

$$v(y) := \overline{F}(y) \left( \int_y^\infty \frac{dF(z)}{\overline{G}(z)} \right)^{-1/2}.$$

*Then*

$$v(y_n)\sqrt{n} \left( \frac{\widehat{F}_N(y_n)}{\overline{F}(y_n)} - 1 \right) = \begin{cases} \xi_n & \text{if } \gamma_F < \gamma_G, \\ O_{\mathbb{P}}(1) & \text{if } \gamma_F = \gamma_G, \end{cases}$$

*where  $\xi_n$  is a random variable which is asymptotically standard Gaussian distributed.*

We now establish the asymptotic normality of  $\widehat{q}_N(\alpha_n)$ .

**Theorem 2** *Let  $\alpha_n \rightarrow 0$ . Assume that  $F$  is a differentiable function in a neighborhood of infinity such that  $yF'(y)/\overline{F}(y) \rightarrow 1/\gamma_F$  as  $y \rightarrow \infty$ , that **(M)** and (3) hold, and that  $nv^2(q(\alpha_n)) \rightarrow \infty$ . Then*

$$v(q(\alpha_n))\sqrt{n} \left( \frac{\widehat{q}_N(\alpha_n)}{q(\alpha_n)} - 1 \right) = \begin{cases} \zeta_n & \text{if } \gamma_F < \gamma_G, \\ O_{\mathbb{P}}(1) & \text{if } \gamma_F = \gamma_G, \end{cases}$$

*where  $\zeta_n$  is a random variable which is asymptotically Gaussian centered with variance  $\gamma_F^2$ .*

Theorem 2 is a convergence result for the intermediate quantile estimator  $\widehat{q}_N(\alpha_n)$ , provided  $nv^2(q(\alpha_n)) \rightarrow \infty$ , which ensures that  $\alpha_n \rightarrow 0$  slowly enough. To examine the asymptotic properties of the extreme quantile estimator (2), we start by proving a couple of results on the tail index estimator  $\widehat{\gamma}_{N,F}(k_N, k'_N)$ . Before that, we introduce some notation: we will write  $\Psi \in \mathcal{RV}_2(a, \Delta)$  where  $a \in \mathbb{R}$  and  $\Delta$  is a bounded measurable function having ultimately constant sign and converging to 0 at infinity, such that  $|\Delta|$  is an ultimately monotonic and regularly varying function, if there exists a positive constant  $c$  such that

$$\Psi(y) = cy^a \exp \left( \int_1^y \frac{\widetilde{\Delta}(z)}{z} dz \right) \text{ with } \widetilde{\Delta}(y) = \Delta(y)(1 + o(1)) \text{ as } y \rightarrow \infty.$$

The following second-order condition is required:

**(C)** We have  $\overline{F} \in \mathcal{RV}_2(-1/\gamma_F, \Delta_F)$  and  $\overline{G} \in \mathcal{RV}_2(-1/\gamma_G, \Delta_G)$  where  $\gamma_F, \gamma_G > 0$  and  $|\Delta_F| \in \mathcal{RV}_1(\rho_F), |\Delta_G| \in \mathcal{RV}_1(\rho_G)$  with  $\rho_F, \rho_G \leq 0$ .

It can be shown (see Lemma 8) that if **(C)** holds, then provided  $\rho_F \neq \rho_G$  and  $\rho_G \neq -1/\gamma_F$ , an analogue of condition **(C)** also holds for  $\overline{F^*}$  and  $\overline{G^*}$ . Finally, let  $U_{F^*}, U_{G^*}$  be the left-continuous inverses of  $1/\overline{F^*}$  and  $1/\overline{G^*}$ . The following result, in which we write  $s \vee t$  and  $s \wedge t$  for the maximum and the minimum of two real numbers  $s$  and  $t$  and  $\lfloor \cdot \rfloor$  for the floor function, examines the asymptotic properties of  $\hat{\gamma}_{N,F}(k_N, k'_N)$ .

**Theorem 3** *Let  $(k_n), (k'_n)$  be such that  $k_n \wedge k'_n \rightarrow \infty$  and  $(k_n \vee k'_n)/n \rightarrow 0$ . Assume that **(M)** holds. Then we have  $\hat{\gamma}_{N,F}(k_N, k'_N) \xrightarrow{\mathbb{P}} \gamma_F$ . Suppose moreover that **(C)** holds, that  $\rho_F \neq \rho_G$  and  $\rho_G \neq -1/\gamma_F$ , that  $k_n \Delta_{F^*}^2(U_{F^*}(n/k_n)) \vee k'_n \Delta_{G^*}^2(U_{G^*}(n/k'_n)) \rightarrow 0$  and*

$$\sup_{r,s \in I_n} \left| \frac{k_r \wedge k'_r}{k_s \wedge k'_s} - 1 \right| \rightarrow 0 \text{ where } I_n = [np(1 - n^{-1/4}), np(1 + n^{-1/4})]. \quad (4)$$

Then if either  $k_n/k'_n \rightarrow 0$  or  $k'_n/k_n \rightarrow 0$ , we have

$$\sqrt{k_{\lfloor np \rfloor} \wedge k'_{\lfloor np \rfloor}} (\hat{\gamma}_{N,F}(k_N, k'_N) - \gamma_F) \xrightarrow{d} \mathcal{N}(0, \sigma_F^2), \quad (5)$$

where  $\sigma_F^2$  is equal to  $\gamma_F^2(1 + \gamma_F/\gamma_G)^2$  if  $k_n/k'_n \rightarrow 0$  and  $\gamma_F^4/\gamma_G^2$  if  $k'_n/k_n \rightarrow 0$ . In the case  $k_n/k'_n \rightarrow 1$ , then we have

$$\sqrt{k_{\lfloor np \rfloor}} (\hat{\gamma}_{N,F}(k_N, k'_N) - \gamma_F) = O_{\mathbb{P}}(1). \quad (6)$$

A careful examination of the proof reveals that contrary to Theorems 1 and 2, Theorem 3 also holds when  $\gamma_F > \gamma_G$ . Before combining Theorems 2 and 3 to obtain the rate of convergence of the estimator  $\hat{q}_N^W(\beta_n | \alpha_n, k_N, k'_N)$ , we state three remarks about Theorem 3.

**Remark 1** *Note that without assuming condition (4) the rate of convergence in Theorem 3 would be the random quantity  $\sqrt{k_N \wedge k'_N}$ , see the proof for further details.*

**Remark 2** *Conditions  $k_n \Delta_{F^*}^2(U_{F^*}(n/k_n)) \rightarrow 0$  and  $k'_n \Delta_{G^*}^2(U_{G^*}(n/k'_n)) \rightarrow 0$  are analogues of the condition classically used to prove the asymptotic normality of the Hill estimator. They ensure that the bias of the estimator is negligible with respect to its standard deviation.*

**Remark 3** *Since in the case  $k_n/k'_n \rightarrow 1$  the correlation between  $\hat{\gamma}_{N,F^*}(k_N)$  and  $\hat{\gamma}_{N,G}(k'_N)$  is very difficult to evaluate, Theorem 3 only provides the rate of convergence of  $\hat{\gamma}_{N,F}(k_N, k'_N) - \gamma_F$  to zero. This case is interesting when  $\gamma_F < \gamma_G$ , i.e. if the tail of  $T$  is larger than the tail of  $Y$ , because then the tail of  $Y$  would not be too contaminated as a result of truncation. On the contrary, in the case when the tail of  $Y$  is the heaviest (which we cannot consider for the estimation of extreme quantiles), there would be higher confidence in the estimation of  $\gamma_G$  than in that of  $\gamma_{F^*}$ , which would lead us to take  $k_n/k'_n \rightarrow 0$ .*

The final result of this section gives some asymptotic properties of the estimator  $\hat{q}_N^W(\beta_n | \alpha_n, k_N, k'_N)$ :

**Theorem 4** *Let  $\alpha_n \rightarrow 0$ ,  $\beta_n \rightarrow 0$ ,  $k_n \wedge k'_n \rightarrow \infty$  and  $(k_n \vee k'_n)/n \rightarrow 0$ . Assume that **(M)**, **(3)** and **(C)** hold. Assume that  $\rho_F \neq \rho_G$  and  $\rho_G \neq -1/\gamma_F$ , that  $\beta_n/\alpha_n \rightarrow 0$ ,  $nv^2(q(\alpha_n)) \rightarrow \infty$ ,  $nv^2(q(\alpha_n))\Delta_F^2(q(\alpha_n)) \rightarrow 0$ ,  $k_n\Delta_{F^*}^2(U_{F^*}(n/k_n)) \vee k'_n\Delta_{G^*}^2(U_{G^*}(n/k'_n)) \rightarrow 0$ ,*

$$(k_{[np]} \wedge k'_{[np]})/nv^2(q(\alpha_n)) \rightarrow 1 \text{ and } \sup_{r,s \in I_n} \left| \frac{k_r \wedge k'_r}{k_s \wedge k'_s} - 1 \right| \rightarrow 0.$$

*Then, if  $\gamma_F < \gamma_G$  and either  $k_n/k'_n \rightarrow 0$  or  $k'_n/k_n \rightarrow 0$ , we have*

$$\frac{v(q(\alpha_n))\sqrt{n}}{\log(\alpha_n/\beta_n)} \left( \frac{\hat{q}_N^W(\beta_n | \alpha_n, k_N, k'_N)}{q(\beta_n)} - 1 \right) \xrightarrow{d} \mathcal{N}(0, \sigma_F^2). \quad (7)$$

*In the case  $k_n/k'_n \rightarrow 1$ , or if  $\gamma_F = \gamma_G$  and either  $k_n/k'_n \rightarrow 0$  or  $k'_n/k_n \rightarrow 0$ , we have*

$$\frac{v(q(\alpha_n))\sqrt{n}}{\log(\alpha_n/\beta_n)} \left( \frac{\hat{q}_N^W(\beta_n | \alpha_n, k_N, k'_N)}{q(\beta_n)} - 1 \right) = O_{\mathbb{P}}(1). \quad (8)$$

## 4 Simulation study

To illustrate the behavior of our estimators, we shall use the following model:

$$\forall y, t > 0, \bar{F}(y) = (1 + y^{1/\delta})^{-\delta/\gamma_F} \text{ and } \bar{G}(t) = (1 + t^{1/\delta})^{-\delta/\gamma_G},$$

where  $\delta > 0$  and  $0 < \gamma_F < \gamma_G$ . Note that in this situation,  $\rho_F = \rho_G = -1/\delta$ . Thus, the larger the value of  $\delta$ , the smaller the values of  $|\rho_F|$  and  $|\rho_G|$  and the slower  $\hat{\gamma}_{N,F}(k_N, k'_N)$  converges to  $\gamma_F$ , see *e.g.* de Haan and Ferreira (2006, p.77) for the related situation when there is no truncation. An important bias in the estimation of  $q(\beta_n)$  can then be expected to appear if  $\delta$  is large. Moreover, the truncation probability is given in our setting by  $1 - p$  with  $p = \gamma_G/(\gamma_F + \gamma_G)$ . In this simulation study, we examine the finite sample behavior of several estimators of the extreme quantile  $q(\beta_n)$  for  $\beta_n$  varying in  $(0, 0.15]$ . We first consider the two estimators introduced in the present paper, namely  $\hat{q}_N(\beta_n)$  and  $\hat{q}_N^W(\beta_n | \alpha_n) = \hat{q}_N(\alpha_n) (\alpha_n/\beta_n)^{\hat{\gamma}_{N,F}(\alpha_n)}$ , where  $(\alpha_n)$  is a sequence in  $(0, 1)$  and  $\hat{\gamma}_{N,F}(\alpha_n)$  is the estimator of the tail-index  $\gamma_F$  defined in (1) with  $k_N = k'_N = \lfloor N\alpha_n \rfloor$ . In order to evaluate how important it is to take into account the fact that the data are right-truncated, we also consider the naive estimators  $\hat{q}_N^*(\beta_n) = \inf\{y \geq y_0 | \hat{\bar{F}}_N^*(y) \leq \beta_n\}$ , and its extrapolated version  $\hat{q}_N^{W,*}(\beta_n) = \hat{q}_N^*(\alpha_n^*) (\alpha_n^*/\beta_n)^{\hat{\gamma}_{N,F^*}(\alpha_n^*)}$ , where  $(\alpha_n^*)$  is a sequence in  $(0, 1)$  and  $\hat{\gamma}_{N,F^*}(\alpha_n^*) = \hat{\gamma}_{N,F^*}(\lfloor N\alpha_n^* \rfloor)$ . These last two estimators are computed using only the observations  $\{Y_i^*, i = 1, \dots, N\}$ , ignoring the fact that these observations are in fact truncated. Note finally that the Weissman-type estimators only depend on the choice of the parameters  $\alpha_n$  or  $\alpha_n^*$ .



In this simulation study, we generate  $R = 1000$  samples of size  $n = 200$  from the distributions  $F$  and  $G$  with  $\delta \in \{1/3, 1\}$ , for  $\gamma_F \in \{1/4, 1/2, 1\}$  and  $p \in \{0.7, 0.8, 0.9, 0.95\}$ . In each case and for given  $\alpha_n, \alpha_n^*$  and  $\beta_n$ , we obtain the observations  $(\hat{q}_N^{(r)}(\beta_n), \hat{q}_N^{W,(r)}(\beta_n|\alpha_n), \hat{q}_N^{(r),*}(\beta_n), \hat{q}_N^{W,(r),*}(\beta_n|\alpha_n^*)), r = 1, \dots, R$  of the estimators. For each replication,  $\alpha_n$  and  $\alpha_n^*$  are then taken as

$$\alpha_{opt}^{(r)} := \arg \min_{\alpha \in (0.04, 0.15]} \int_{0.04}^{0.15} \log^2 \left( \frac{\hat{q}_N^{(r)}(\beta)}{\hat{q}_N^{W,(r)}(\beta|\alpha)} \right) d\beta,$$

$$\text{and } \alpha_{opt}^{(r),*} := \arg \min_{\alpha \in (0.04, 0.15]} \int_{0.04}^{0.15} \log^2 \left( \frac{\hat{q}_N^{(r),*}(\beta)}{\hat{q}_N^{W,(r),*}(\beta|\alpha)} \right) d\beta.$$

The idea behind these criteria is that for quantiles which are not too large, the estimators  $\hat{q}_N^{(r)}$  (resp.  $\hat{q}_N^{(r),*}$ ) and  $\hat{q}_N^{W,(r)}(\cdot|\alpha_n)$  (resp.  $\hat{q}_N^{W,(r),*}(\cdot|\alpha_n^*)$ ) should be close if  $\alpha_n$  (resp.  $\alpha_n^*$ ) is well chosen. Next, we compute the errors

$$E(\check{q}^{(r)}) := \int_0^{0.15} \log^2 \left( \frac{\check{q}^{(r)}(\beta)}{q(\beta)} \right) d\beta,$$

where  $\check{q}^{(r)}$  is either  $\hat{q}_N^{(r)}, \hat{q}_N^{(r),*}, \hat{q}_N^{W,(r)}(\cdot|\alpha_{opt}^{(r)})$  or  $\hat{q}_N^{W,(r),*}(\cdot|\alpha_{opt}^{(r),*})$ . The error  $E$  is a measure of the overall performance of a quantile estimator when estimating extreme quantiles. For  $\theta = \{0.1, 0.5, 0.9\}$ , let  $r(\theta)$  (resp.  $s(\theta), r^*(\theta)$  and  $s^*(\theta)$ ) be the replication corresponding to the quantile of order  $\theta$  of the set  $\{E(\hat{q}_N^{(r)}), r = 1, \dots, R\}$  (resp.  $\{E(\hat{q}_N^{W,(r)}), r = 1, \dots, R\}, \{E(\hat{q}_N^{(r),*}), r = 1, \dots, R\}$  and  $\{E(\hat{q}_N^{W,(r),*}), r = 1, \dots, R\}$ ). In each situation, the errors corresponding to these replications are given respectively in Tables 1, 2, 3 and 4. It appears that the Weissman-type estimators perform better than the others, which is not surprising since these estimators are specifically adapted to the estimation of extreme quantiles. Besides, we can see that the smaller is the probability  $p$ , the larger is the bias of the estimators. This was also expected since in our setting, a smaller probability  $p$  means a greater number of observations missing in the right tail of  $Y$ . Note also that the importance to take into account the fact that the data are right-truncated appears clearly for the Weissman-type estimators when  $p$  is small. For large values of  $p$ , the naive estimators and the estimators proposed in this paper have similar performances.

## 5 Real data example

The dataset we work on here deals with the lifetime of automobile brake pads. It was already considered by Lawless (2002, Example 2.4.2) and obtained in the following way: in order to study the brake pad lifetime, a car manufacturer selected a random sample of cars which were sold over the previous year. In this situation, one way to obtain the brake pad lifetime is to conduct frequent assessments of the brake pad until it is found to be so worn out that it requires replacement, but this procedure is too time-consuming and difficult to

implement. Instead, this lifetime was estimated by the manufacturer, based on a measure of the brake pad thickness. For each car, the observed mileage  $M$  and the estimated lifetime  $L$  were collected, and only the cars with a response variable  $L$  larger than  $M$ , or equivalently cars whose brake pad thickness was so large that the brake pads did not require immediate replacement, were kept. The random variables  $L$  and  $M$  can reasonably be assumed to be independent. At the end of this process, only  $N = 98$  estimated lifetimes  $\{L_i^*, i = 1, \dots, N\}$  and mileages  $\{M_i^*, i = 1, \dots, N\}$  remained in the sample. Note that since the variable of interest is the brake pad lifetime, this is actually a randomly left-truncated sample. In order to work within the framework of the present paper, the following transformation is used: for  $i \in \{1, \dots, N\}$ , we define

$$Y_i^* = (L_i^* - m_N - \varepsilon t)^{-1} \text{ and } T_i^* = (M_i^* - m_N - \varepsilon)^{-1},$$

where  $m_N = \min\{M_1^*, \dots, M_N^*\}$  and  $\varepsilon = 0.05$ . Thus,  $(Y_1^*, T_1^*), \dots, (Y_N^*, T_N^*)$  can be seen as randomly right-truncated observations from a random sample of independent copies of a random pair  $(Y, T)$ , whose size  $n$  is unknown. We now need to check that this random sample presents some evidence of heavy tails. To this aim, using our Hill-type estimators, the estimations of the tail indices  $\gamma_{F^*}$  and  $\gamma_G$  (using the same notation as before) are represented on Figure 7 as functions of  $k_N \in \{1, \dots, 90\}$ . It appears clearly that the estimated values are positive and, for intermediate values of  $k_N$ , the estimations of  $\gamma_{F^*}$  and  $\gamma_G$  seem to be fairly stable. We thus are led to believe that there is indeed evidence of heavy tails for this random sample.

We now estimate the extreme quantiles of  $Y$  using the Weissman-type estimator  $\hat{q}_N(\beta_n|\alpha_n)$ , where  $\alpha_n$  is chosen as in the simulation study. The selected value of  $\alpha_n$  for this dataset is 0.123, corresponding to  $k_N = 12$ . The tail indices  $\gamma_{F^*}$ ,  $\gamma_F$  and  $\gamma_G$  are respectively estimated to be 0.32, 0.60 and 0.69. In particular, the estimate of  $\gamma_F$  is less than that of  $\gamma_G$ , which seems to indicate that condition **(M)** is satisfied. On Figure 7, the estimated quantile  $\hat{q}_N(\beta_n|\alpha_n)$  for  $\beta_n \in (0, 0.15)$  and the one associated to the original data, namely  $\tilde{q}_N(1 - \beta_n|\alpha_n) = (\hat{q}_N(\beta_n|\alpha_n))^{-1} + m_N - \varepsilon$  are represented. In particular, we may see on the right panel of Figure 7 that the brake pad lifetime is estimated to be less than 13 500 kilometers for only 1% of the cars.

## 6 Conclusion

In this paper, we introduced and studied an extreme quantile estimator for randomly right-truncated data. This estimator is built upon an empirical high quantile estimator and a tail index estimator which are both adapted for random right-truncation. Our extreme quantile estimator has satisfying performances, be them theoretical or practical.

Future work on our estimator includes obtaining the asymptotic normality of the tail index estimator and the extreme quantile estimator when  $k_n = k'_n$ . This is not only a stimulating theoretical problem but also an interesting practical one,

since we consider this case in our simulation study and data analysis. Another question worthy of research would be to build and study extreme-value index estimators and extreme quantile estimators when the distribution functions of  $Y$  and  $T$  belong to an arbitrary max-domain of attraction (de Haan and Ferreira 2006). This would certainly be useful in practical applications, for instance when trying to handle random samples whose underlying marginal distributions are believed to be short-tailed.

## 7 Proofs of the main results

**Proof of Theorem 1.** As a preliminary step, note that when  $N > 0$  (which is true with arbitrarily large probability as  $n \rightarrow \infty$ , by Lemma 1), one has

$$\begin{aligned} \frac{\hat{\Lambda}_N^F(y_n) - \Lambda^F(y_n)}{\bar{F}(y_n)} &= S_{n,1} + S_{n,2} + S_{n,3} - \frac{\Lambda^F(y_n)}{\bar{F}(y_n)} \mathbb{I}_{\{Y_{N,N}^* \leq y_n\}} \\ \text{where } S_{n,1} &= \mathbb{I}_{\{Y_{N,N}^* > y_n\}} \left( \frac{1}{n} \left[ \sum_{i=1}^N \frac{\mathbb{I}_{\{Y_i^* > y_n\}}}{p\bar{F}(y_n)C^*(Y_i^*)} \right] - \frac{\Lambda^F(y_n)}{\bar{F}(y_n)} \right), \\ S_{n,2} &= \left( \frac{p}{N} - \frac{1}{n} \right) \left( \sum_{i=1}^N \frac{\mathbb{I}_{\{Y_i^* > y_n\}}}{p\bar{F}(y_n)C^*(Y_i^*)} \right) \mathbb{I}_{\{Y_{N,N}^* > y_n\}} \\ \text{and } S_{n,3} &= \frac{\mathbb{I}_{\{Y_{N,N}^* > y_n\}}}{N\bar{F}(y_n)} \sum_{i=1}^N \mathbb{I}_{\{Y_i^* > y_n\}} \left( \frac{1}{\hat{C}_N^*(Y_i^*)} - \frac{1}{C^*(Y_i^*)} \right). \end{aligned}$$

Lemma 4 entails that  $\mathbb{I}_{\{Y_{N,N}^* \leq y_n\}}$  is zero with arbitrarily large probability as  $n \rightarrow \infty$  and thus:

$$v(y_n)\sqrt{n} \left( \frac{\hat{\Lambda}_N^F(y_n) - \Lambda^F(y_n)}{\bar{F}(y_n)} \right) = v(y_n)\sqrt{n} \sum_{j=1}^3 S_{n,j} + o_{\mathbb{P}}(1).$$

Let us first focus on the term  $S_{n,1}$  which we can rewrite as

$$S_{n,1} = \frac{\mathbb{I}_{\{Y_{N,N}^* > y_n\}}}{n} \sum_{i=1}^n W_{n,i} \text{ where } W_{n,i} := \frac{\mathbb{I}_{\{Y_i > y_n\}} \mathbb{I}_{\{Y_i \leq T_i\}}}{p\bar{F}(y_n)C^*(Y_i)} - \frac{\Lambda^F(y_n)}{\bar{F}(y_n)}. \quad (9)$$

It is easy to check that the  $W_{n,i}$  are independent, identically distributed and centered random variables. From (9) we get

$$\mathbb{E}(W_{n,1}^2) = \frac{1}{p^2 \bar{F}^2(y_n)} \mathbb{E} \left( \frac{\mathbb{I}_{\{Y > y_n\}} \mathbb{I}_{\{Y \leq T\}}}{(C^*(Y))^2} \right)^2 - \left( \frac{\Lambda^F(y_n)}{\bar{F}(y_n)} \right)^2. \quad (10)$$

We now use the fact that since  $F$  is nondecreasing and  $F(y_n) \rightarrow 1$ , we have

$$\frac{\Lambda^F(y_n)}{\bar{F}(y_n)} - 1 = \frac{1}{\bar{F}(y_n)} \int_{y_n}^{\infty} \frac{\bar{F}(z)}{F(z)} dF(z) = O(\bar{F}(y_n)) \quad (11)$$

and by Proposition 1,  $C^* = p^{-1}F\overline{G}$  so that

$$\frac{1}{p^2}\mathbb{E}\left(\frac{\mathbb{I}_{\{Y>y_n\}}\mathbb{I}_{\{Y\leq T\}}}{(C^*(Y))^2}\right) = \int_{y_n}^{\infty} \frac{dF(z)}{\overline{G}(z)F^2(z)} = \int_{y_n}^{\infty} \frac{dF(z)}{\overline{G}(z)}(1 + o(1)). \quad (12)$$

• In the case  $\gamma_F < \gamma_G$ , noting that  $dF(z) = -d\overline{F}(z)$ , we get from (10), (11), (12) and (36) (see the Appendix):

$$\mathbb{E}(W_{n,1}^2) = \frac{\gamma_G}{\gamma_G - \gamma_F} \frac{1}{\overline{F}(y_n)\overline{G}(y_n)}(1 + o(1)). \quad (13)$$

Furthermore, because  $\gamma_F < \gamma_G$ , one can pick  $\delta > 0$  such that  $(1+\delta)/\gamma_G - 1/\gamma_F < 0$ . Hölder's inequality then gives

$$\mathbb{E}[|W_{n,1}|^{2+\delta}] \leq 2^{1+\delta} \left( \mathbb{E} \left( \frac{\mathbb{I}_{\{Y_i>y_n\}}\mathbb{I}_{\{Y_i\leq T_i\}}}{p\overline{F}(y_n)C^*(Y_i)} \right)^{2+\delta} + 1 + o(1) \right),$$

where (11) was used. Besides, since  $F$  is nondecreasing and  $F(y_n) \rightarrow 1$ , we obtain by Lemma 2 with  $\varphi = 1/\overline{G}^{1+\delta}$  and  $\psi = \overline{F}$ :

$$\mathbb{E} \left( \frac{\mathbb{I}_{\{Y_i>y_n\}}\mathbb{I}_{\{Y_i\leq T_i\}}}{pC^*(Y_i)} \right)^{2+\delta} = \int_{y_n}^{\infty} \frac{dF(z)}{\overline{G}^{1+\delta}(z)F^{2+\delta}(z)} = O \left( \frac{\overline{F}(y_n)}{\overline{G}^{1+\delta}(y_n)} \right). \quad (14)$$

It follows from (13), (14) and (36) that

$$n^{-\delta/2} \frac{\mathbb{E}[|W_{n,1}|^{2+\delta}]}{[\text{Var}(W_{n,1})]^{1+\delta/2}} = O([n\overline{F}(y_n)\overline{G}(y_n)]^{-\delta}) = O([v(y_n)\sqrt{n}]^{-\delta}) \rightarrow 0.$$

Since the  $W_{n,i}$  are independent, identically distributed and centered random variables, Lyapunov's central limit theorem (see e.g. Billingsley 1979, p.312) entails  $\sqrt{n}S_{n,1}/\sqrt{\text{Var}(W_{n,1})} \xrightarrow{d} \mathcal{N}(0,1)$ . Using (13) and the convergence  $\mathbb{I}_{\{Y_{N,N}^*>y_n\}} \rightarrow 1$  leads to

$$v(y_n)\sqrt{n}S_{n,1} \xrightarrow{d} \mathcal{N}(0,1) \text{ when } \gamma_F < \gamma_G. \quad (15)$$

• When  $\gamma_F = \gamma_G$ , we note that using condition (3), the second part of Lemma 2 with  $\varphi = 1/\overline{G}$  and  $\psi = \overline{F}$  entails that the function  $v$  is regularly varying with index  $-1/\gamma_F < 0$ . This yields

$$\frac{1}{\overline{F}^2(y_n)} \int_{y_n}^{\infty} \frac{dF(z)}{\overline{G}(z)} \rightarrow \infty. \quad (16)$$

Consequently, from (10), (11) and (16), we get  $\mathbb{E}(W_{n,1}^2) = O(1/v^2(y_n))$ , which entails

$$v(y_n)\sqrt{n}S_{n,1} = O_{\mathbb{P}}(1) \text{ when } \gamma_F = \gamma_G. \quad (17)$$

Let us now focus on the term  $S_{n,2}$ . From the previous results, it is clear that

$$\frac{1}{n} \sum_{i=1}^N \frac{\mathbb{I}_{\{Y_i^* > y_n\}}}{p\bar{F}(y_n)C^*(Y_i^*)} \mathbb{I}_{\{Y_{N,N}^* > y_n\}} \xrightarrow{\mathbb{P}} 1.$$

Since  $np/N = 1 + O_{\mathbb{P}}(n^{-1/2})$  from Lemma 1, one has, using Lemma 4, that  $S_{n,2} = O_{\mathbb{P}}(n^{-1/2})$ . Using the convergence  $v(y_n) \rightarrow 0$  it is now obvious that

$$v(y_n)\sqrt{n}S_{n,2} = o_{\mathbb{P}}(1). \quad (18)$$

Let us now control  $S_{n,3}$ . Note that (see Woodroffe 1985, pp.172–173):

$$\sqrt{n} \sup_{z \in \mathbb{R}} \left| \hat{F}_N^*(z) - F^*(z) \right| = O_{\mathbb{P}}(1) \text{ and } \sqrt{n} \sup_{z \in \mathbb{R}} \left| \hat{G}_N^*(z) - G^*(z) \right| = O_{\mathbb{P}}(1).$$

Therefore, it comes that

$$\sqrt{n} \sup_{1 \leq i \leq N} \left| \hat{C}_N^*(Y_i^*) - C^*(Y_i^*) \right| \mathbb{I}_{\{Y_i^* > y_n\}} = O_{\mathbb{P}}(1).$$

By Lemmas 1 and 5, we thus obtain:

$$v(y_n)\sqrt{n}S_{n,3} = O_{\mathbb{P}} \left( \frac{v(y_n)}{n\bar{F}(y_n)} \sum_{i=1}^N \frac{\mathbb{I}_{\{Y_i^* > y_n\}}}{(C^*(Y_i^*))^2} \mathbb{I}_{\{Y_{N,N}^* > y_n\}} \right). \quad (19)$$

Since

$$\sum_{i=1}^N \frac{\mathbb{I}_{\{Y_i^* > y_n\}}}{(C^*(Y_i^*))^2} = \sum_{i=1}^n \frac{\mathbb{I}_{\{Y_i > y_n\}} \mathbb{I}_{\{Y_i \leq T_i\}}}{(C^*(Y_i))^2},$$

it follows from (12) that

$$\mathbb{E} \left( \frac{v(y_n)}{n\bar{F}(y_n)} \sum_{i=1}^N \frac{\mathbb{I}_{\{Y_i^* > y_n\}}}{(C^*(Y_i^*))^2} \right) = p^2 \sqrt{\int_{y_n}^{\infty} \frac{dF(z)}{\bar{G}(z)}} (1 + o(1)) \rightarrow 0, \quad (20)$$

because the integral in the right-hand side converges to 0. Using (19) and (20) entails

$$v(y_n)\sqrt{n}S_{n,3} = O_{\mathbb{P}} \left( \frac{v(y_n)}{n\bar{F}(y_n)} \sum_{i=1}^N \frac{\mathbb{I}_{\{Y_i^* > y_n\}}}{(C^*(Y_i^*))^2} \mathbb{I}_{\{Y_{N,N}^* \geq y_n\}} \right) = o_{\mathbb{P}}(1). \quad (21)$$

Use finally (15), (17), (18) and (21) together to get

$$v(y_n)\sqrt{n} \left( \frac{\hat{\Lambda}_N^F(y_n) - \Lambda^F(y_n)}{\bar{F}(y_n)} \right) = \begin{cases} \xi_n & \text{if } \gamma_F < \gamma_G, \\ O_{\mathbb{P}}(1) & \text{if } \gamma_F = \gamma_G, \end{cases}$$

where  $\xi_n$  is a random variable which is asymptotically standard Gaussian distributed. Using the delta-method concludes the proof of Theorem 1.  $\blacksquare$

**Proof of Theorem 2.** We start by the case  $\gamma_F < \gamma_G$  and we define  $\sigma_n = q(\alpha_n)/[v(q(\alpha_n))\sqrt{n}]$ . It is enough to show that  $\Phi_n(z) := \mathbb{P}(\sigma_n^{-1}(\widehat{q}_N(\alpha_n) - q(\alpha_n)) \leq z) \rightarrow \Phi(z)$ , for every  $z \in \mathbb{R}$ , where  $\Phi$  is the cdf of a  $\mathcal{N}(0, \gamma_F^2)$  distribution. Let us introduce the sequence  $\vartheta_n := \gamma_F v(q(\alpha_n))\sqrt{n}/\alpha_n$ . It is easy to check that  $\Phi_n(z) = \mathbb{P}(W_n \leq z_n)$ , where

$$W_n = \vartheta_n \left( \widehat{F}_N(\widetilde{q}_n) - \overline{F}(\widetilde{q}_n) \right) \text{ and } z_n = \vartheta_n(\alpha_n - \overline{F}(\widetilde{q}_n)),$$

with  $\widetilde{q}_n = q(\alpha_n) + \sigma_n z$ . Let us first focus on the nonrandom term  $z_n$ . Since  $F$  is a differentiable function, there exists  $\theta_n \in (0, 1)$  such that  $\alpha_n - \overline{F}(\widetilde{q}_n) = \sigma_n z F'(q(\alpha_n) + \theta_n \sigma_n z)$ . Since  $\sigma_n/q(\alpha_n) \rightarrow 0$  as  $n \rightarrow \infty$ , we may use the convergence  $yF'(y)/\overline{F}(y) \rightarrow 1/\gamma_F$  to get  $F'(q(\alpha_n) + \theta_n \sigma_n z) = \gamma_F^{-1} \alpha_n/q(\alpha_n)(1 + o(1))$ . Hence the following equality holds:

$$z_n = \vartheta_n \sigma_n z \frac{1}{\gamma_F} \frac{\alpha_n}{q(\alpha_n)} (1 + o(1)) = z(1 + o(1)). \quad (22)$$

We now consider the random term  $W_n$ . One has

$$W_n = \frac{\vartheta_n \overline{F}(\widetilde{q}_n)}{v(\widetilde{q}_n)\sqrt{n}} Z_n \text{ where } Z_n = v(\widetilde{q}_n)\sqrt{n} \left( \frac{\widehat{F}_N(\widetilde{q}_n)}{\overline{F}(\widetilde{q}_n)} - 1 \right).$$

Note that from model **(M)**,  $\overline{F}(\widetilde{q}_n) = \alpha_n(1 + o(1))$ . Moreover, it is a consequence of Lemma 2 that the function  $v$  is regularly varying, so that  $v(\widetilde{q}_n) = v(q(\alpha_n))(1 + o(1))$ . Consequently  $\vartheta_n \overline{F}(\widetilde{q}_n) = \gamma_F v(\widetilde{q}_n)\sqrt{n}(1 + o(1))$ . Apply then Theorem 1 with  $y_n = \widetilde{q}_n$  to obtain that  $Z_n \xrightarrow{d} \mathcal{N}(0, 1)$  and thus  $W_n \xrightarrow{d} \mathcal{N}(0, \gamma_F^2)$ , which concludes the proof in the case  $\gamma_F < \gamma_G$ .

Now, if  $\gamma_F = \gamma_G$ , we start by showing that if  $(\varepsilon_n)$  is an arbitrary nonrandom positive sequence tending to 0 at infinity such that  $\varepsilon_n v(q(\alpha_n))\sqrt{n} = \varepsilon_n q(\alpha_n)/\sigma_n \rightarrow \infty$ , we have

$$\varepsilon_n \sigma_n^{-1} |\widehat{q}_N(\alpha_n) - q(\alpha_n)| \xrightarrow{\mathbb{P}} 0. \quad (23)$$

Pick then an arbitrary  $z > 0$ . We shall show that  $\varphi_n(z) := \mathbb{P}(\varepsilon_n \sigma_n^{-1} |\widehat{q}_N(\alpha_n) - q(\alpha_n)| > z) \rightarrow 0$ . With  $\vartheta_n$  as above, it is easy to check that  $\varphi_n(z) = \mathbb{P}(W_{n,+} > z_{n,+}) + \mathbb{P}(W_{n,-} < z_{n,-})$ , where

$$W_{n,\pm} = \vartheta_n \left( \widehat{F}_N(\widetilde{q}_n) - \overline{F}(\widetilde{q}_n) \right) \text{ and } z_{n,\pm} = \vartheta_n(\alpha_n - \overline{F}(\widetilde{q}_n)),$$

where we redefine  $\widetilde{q}_n := q(\alpha_n) \pm \varepsilon_n^{-1} \sigma_n z$ . Let us first focus on the nonrandom term  $z_{n,\pm}$ . Mimicking the arguments leading to (22) in the proof of the first part of Theorem 2, we get that

$$z_{n,\pm} = \pm \vartheta_n \varepsilon_n^{-1} \sigma_n z \frac{1}{\gamma_F} \frac{\alpha_n}{q(\alpha_n)} (1 + o(1)) = \pm \varepsilon_n^{-1} z (1 + o(1)). \quad (24)$$

We now consider the random term  $W_{n,\pm}$ . One has

$$W_{n,\pm} = \frac{\vartheta_n \overline{F}(\widetilde{q}_n)}{v(\widetilde{q}_n)\sqrt{n}} Z_{n,\pm} \text{ where } Z_{n,\pm} = v(\widetilde{q}_n)\sqrt{n} \left( \frac{\widehat{F}_N(\widetilde{q}_n)}{\overline{F}(\widetilde{q}_n)} - 1 \right).$$

Since  $\bar{F}$  and  $v$  are regularly varying, we have  $\bar{F}(\tilde{q}_n) = \alpha_n(1 + o(1))$  and  $v(\tilde{q}_n) = v(q(\alpha_n))(1 + o(1))$  which leads to  $\vartheta_n \bar{F}(\tilde{q}_n) = \gamma_F v(\tilde{q}_n) \sqrt{n}(1 + o(1))$ . On the other hand, the second part of Theorem 1 implies that  $\varepsilon_n Z_{n,\pm} = o_{\mathbb{P}}(1)$ , so that using (24) we obtain for  $n$  large enough

$$\varphi_n(z) \leq \mathbb{P}(\varepsilon_n Z_{n,+} > z/2) + \mathbb{P}(\varepsilon_n Z_{n,-} < -z/2) \rightarrow 0$$

and the proof of (23) is complete. Note now that if  $(\varepsilon_n)$  is an arbitrary non-random positive sequence, we have  $\varepsilon_n \leq \varepsilon'_n := \varepsilon_n \vee (v(q(\alpha_n))\sqrt{n})^{-1/2}$  with  $\varepsilon'_n v(q(\alpha_n))\sqrt{n} \rightarrow \infty$ . It can thus easily be seen that in fact (23) holds for every positive sequence  $(\varepsilon_n)$ ; applying Lemma 6 completes the proof of Theorem 2. ■

**Proof of Theorem 3.** Use Lemma 1 to get  $\mathbb{P}(N \in I_n) \rightarrow 1$ , where  $I_n$  is defined in equation (4). The consistency statement is thus an immediate consequence of Lemmas 3 and 9.

To prove (5) and (6), write

$$\hat{\gamma}_{n,F}(k_N, k'_N) - \gamma_F = \frac{\hat{\gamma}_{N,F^*}(k_N) \hat{\gamma}_{N,G}(k'_N)}{\hat{\gamma}_{N,G}(k'_N) - \hat{\gamma}_{N,F^*}(k_N)} - \frac{\gamma_{F^*} \gamma_G}{\gamma_G - \gamma_{F^*}}.$$

Since  $\hat{\gamma}_{N,F^*}(k_N) \xrightarrow{\mathbb{P}} \gamma_{F^*}$  and  $\hat{\gamma}_{N,G}(k'_N) \xrightarrow{\mathbb{P}} \gamma_G$ , it is straightforward to obtain the equality

$$\begin{aligned} \hat{\gamma}_{n,F}(k_N, k'_N) - \gamma_F &= \left(1 + \frac{\gamma_F}{\gamma_G}\right)^2 (\hat{\gamma}_{N,F^*}(k_N) - \gamma_{F^*}) - \frac{\gamma_F^2}{\gamma_G^2} (\hat{\gamma}_{N,G}(k'_N) - \gamma_G) \\ &\quad + o_{\mathbb{P}}(\hat{\gamma}_{N,F^*}(k_N) - \gamma_{F^*}) + o_{\mathbb{P}}(\hat{\gamma}_{N,G}(k'_N) - \gamma_G). \end{aligned}$$

Applying Lemma 8 proves that  $\bar{F}^*$  and  $\bar{G}^*$  satisfy the conditions of Lemma 9. Using then Lemma 9 twice concludes the proof. ■

**Proof of Theorem 4.** From condition (C), note that we may write

$$\forall y > 0, \bar{F}(y) = y^{-1/\gamma_F} L_F(y) \text{ with } L_F(y) = c_F \exp \left( \int_1^y \frac{\tilde{\Delta}_F(v)}{v} dv \right)$$

where  $c_F$  is a positive constant and  $\tilde{\Delta}_F$  is asymptotically equivalent to  $\Delta_F$ . Further, since  $q$  is the (generalized) inverse function of  $\bar{F}$ , it satisfies the equation

$$\forall \alpha \in (0, 1), q(\alpha) = \alpha^{-\gamma_F} L_F^{\gamma_F}(q(\alpha)). \quad (25)$$

Note that since  $|\tilde{\Delta}_F|$  is asymptotically equivalent to the ultimately monotonic function  $|\Delta_F|$ , one has for  $n$  large enough

$$\varepsilon_n := \frac{1}{\log(\alpha_n/\beta_n)} \left| \log \left( \frac{L_F(q(\beta_n))}{L_F(q(\alpha_n))} \right) \right| \leq 2 \frac{|\Delta_F(q(\alpha_n))|}{\log(\alpha_n/\beta_n)} \log \left( \frac{q(\beta_n)}{q(\alpha_n)} \right).$$

By (25) we obtain for  $n$  large enough  $\varepsilon_n \leq 2\gamma_F |\Delta_F(q(\alpha_n))| (1 + \varepsilon_n)$ , which entails

$$\varepsilon_n \leq \frac{2\gamma_F |\Delta_F(q(\alpha_n))|}{1 - 2\gamma_F |\Delta_F(q(\alpha_n))|} = O(|\Delta_F(q(\alpha_n))|).$$

Using once again (25), we get

$$\begin{aligned} \log \left( \frac{\widehat{q}_N^W(\beta_n | \alpha_n, k_N, k'_N)}{q(\beta_n)} \right) &= \log \left( \frac{\widehat{q}_N(\alpha_n)}{q(\alpha_n)} \right) + (\widehat{\gamma}_{N,F}(k_N, k'_N) - \gamma_F) \log \left( \frac{\alpha_n}{\beta_n} \right) \\ &\quad + O(|\Delta_F(q(\alpha_n))|). \end{aligned} \quad (26)$$

To prove (7), remark that since  $\log(\alpha_n/\beta_n) \rightarrow \infty$ , applying Theorems 2 and 3 together with Slutsky's lemma yields, if either  $k_n/k'_n \rightarrow 0$  or  $k'_n/k_n \rightarrow 0$  with  $\gamma_F < \gamma_G$ :

$$\frac{v(q(\alpha_n))\sqrt{n}}{\log(\alpha_n/\beta_n)} \log \left( \frac{\widehat{q}_N^W(\beta_n | \alpha_n, k_N, k'_N)}{q(\beta_n)} \right) \xrightarrow{d} \mathcal{N}(0, \sigma_F^2).$$

Using the delta-method ends the proof of (7). To prove (8) if  $k_n/k'_n \rightarrow 1$  or  $\gamma_F = \gamma_G$ , use (26), Theorems 2 and 3 to get

$$\frac{v(q(\alpha_n))\sqrt{n}}{\log(\alpha_n/\beta_n)} \log \left( \frac{\widehat{q}_N^W(\beta_n | \alpha_n, k_N, k'_N)}{q(\beta_n)} \right) = O_{\mathbb{P}}(1).$$

Applying the mean-value theorem to the exponential function ends the proof of Theorem 4. ■

## References

- [1] Beirlant J, Goegebeur Y, Segers J, Teugels J (2004) Statistics of Extremes. John Wiley & Sons, New York
- [2] Beirlant J, Guillou A (2001) Pareto index estimation under moderate right censoring. Scand Actuar J 2:111–125
- [3] 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
- [4] Beirlant J, Guillou A, Toulemonde G (2010) Peaks-over-threshold modeling under random censoring. Comm Statist Theory Methods 39:1158–1179
- [5] Billingsley P (1979) Probability and measure. Wiley, New York
- [6] Bingham NH, Goldie CM, Teugels JL (1987) Regular variation. Cambridge University Press
- [7] Einmahl JHJ, Fils-Villetard A, Guillou A (2008) Statistics of extremes under random censoring. Bernoulli 14(1):207–227
- [8] El Methni J, Gardes L, Girard S (2014) Nonparametric estimation of extreme risks from conditional heavy-tailed distributions. Scand J Stat, to appear



- [9] Gomes MI, Neves MM (2011) Estimation of the extreme value index for randomly censored data. *Biometrical Letters* 48:1–22
- [10] de Haan L, Ferreira A (2006) *Extreme value theory: An introduction*. Springer, New York
- [11] Herbst T (1999) An application of randomly truncated data models in reserving IBNR claims. *Insurance Math Econom* 25(2):123–131
- [12] Hill BM (1975) A simple general approach to inference about the tail of a distribution. *Ann Statist* 3:1163–1174
- [13] Hu XJ, Lawless JF (1996a) Estimation from truncated life time data with supplementary information on covariates and censoring times. *Biometrika* 83(4):747–761
- [14] Hu XJ, Lawless JF (1996b) Estimation of rate and mean functions from truncated recurrent event data. *J Amer Statist Assoc* 91(433):300–310
- [15] Jackson JC (1974) The analysis of quasar samples. *Mon Not Roy Astron Soc* 166:281–295
- [16] Kalbfleisch JD, Lawless JF (1989) Inference based on retrospective ascertainment: an analysis of the data on transfusion-related AIDS. *J Amer Statist Assoc* 84(406):360–372
- [17] Kalbfleisch JD, Lawless JF (1991) Regression models for right truncated data with applications to AIDS incubation times and reporting lags. *Statist Sinica* 1:19–32
- [18] Kaminsky KS (1987) Prediction of IBNR claim counts by modelling the distribution of report lags. *Insurance Math Econom* 6(2):151–159
- [19] Lagakos SW, Barraj LM, De Gruttola V (1988) Nonparametric analysis of truncated survival data, with application to AIDS. *Biometrika* 75(3):515–523
- [20] Lawless JF (1994) Adjustments for reporting delays and the prediction of occurred but not reported events. *Canad J Statist* 22(1):15–31
- [21] Lawless JF (2002) *Statistical models and methods for lifetime data*, 2nd edn. Wiley Series in Probability and Statistics
- [22] Lynden-Bell D (1971) A method of allowing for known observational selection in small samples applied to 3CR quasars. *Mon Not Roy Astron Soc* 155:95–118
- [23] Meeker WQ, Escobar LA (1998) *Statistical methods for reliability data*. John Wiley & Sons, New York
- [24] Resnick S (2007) *Extreme Values, Regular Variation, and Point Processes*. Springer, New York

- [25] Shorack, G., Wellner, J. (1986). *Empirical processes with applications to statistics*. John Wiley and Sons.
- [26] Stute W (1993) Almost sure representations of the product-limit estimator for truncated data. *Ann Statist* 21(1):146–156
- [27] Stute W, Wang JL (2008) The central limit theorem under random truncation. *Bernoulli* 14(3):604–622
- [28] Weissman I (1978) Estimation of parameters and large quantiles based on the  $k$  largest observations. *J Amer Statist Assoc* 73:812–815
- [29] Woodroffe M (1985) Estimating a distribution function with truncated data. *Ann Statist* 13(1):163–177
- [30] Worms J, Worms R (2014) New estimators of the extreme value index under random right censoring, for heavy-tailed distributions. *Extremes* 17:337–358

## Appendix - Preliminary results

The first result gives an equivalent of the random variable  $N$  and the conditional distribution of  $(Y_1^*, T_1^*), \dots, (Y_N^*, T_N^*)$  given  $N$ .

**Lemma 1** *We have  $N/np = 1 + O_{\mathbb{P}}(n^{-1/2})$ . Furthermore, the conditional distribution of  $(Y_1^*, T_1^*), \dots, (Y_N^*, T_N^*)$  given  $N = m > 0$  is equal to the distribution of  $m$  independent copies of a random vector  $(Y^*, T^*)$  with cdf  $H^*$ .*

**Proof of Lemma 1.** The first part of the result is a straightforward consequence of Chebyshev's inequality. Let now  $(A_i)_{i \geq 1}$  be arbitrary Borel subsets of  $[y_0, \infty) \times [t_0, \infty)$ . If  $m \geq 1$ ,

$$\begin{aligned}
& \mathbb{P}((Y_1^*, T_1^*) \in A_1, \dots, (Y_N^*, T_N^*) \in A_N, N = m) \\
&= \binom{n}{m} \mathbb{P}((Y_i, T_i) \in A_i, Y_i \leq T_i, Y_j > T_j, i = 1, \dots, m, j = m+1, \dots, n) \\
&= \left( \binom{n}{m} \prod_{i=1}^m \mathbb{P}(Y_i \leq T_i) \prod_{j=m+1}^n \mathbb{P}(Y_j > T_j) \right) \prod_{i=1}^m \mathbb{P}((Y_i, T_i) \in A_i | Y_i \leq T_i) \\
&= \mathbb{P}(N = m) \prod_{i=1}^m \mathbb{P}((Y^*, T^*) \in A_i),
\end{aligned}$$

which concludes the proof. ■

**Lemma 2** *Let  $\varphi \in \mathcal{RV}_1(\alpha)$  and  $\psi \in \mathcal{RV}_1(-\beta)$  with  $\alpha \in \mathbb{R}$  and  $\beta > 0$ . Assume that  $\psi$  is right-continuous and nonincreasing on some interval  $[A, \infty)$ ,  $A \geq 0$ .*

- If  $\alpha < \beta$  then the function  $\varphi$  is integrable with respect to  $\psi$  on a neighborhood of infinity and

$$\int_y^\infty \varphi(z) d\psi(z) = -\frac{\beta}{\beta - \alpha} \varphi(y) \psi(y) (1 + o(1)) \text{ as } y \rightarrow \infty.$$

- If  $\alpha = \beta$  and  $\int_A^\infty \varphi(z) d\psi(z) < \infty$  then  $y \mapsto -\int_y^\infty \varphi(z) d\psi(z)$  is slowly varying at infinity and

$$-\frac{1}{\varphi(y) \psi(y)} \int_y^\infty \varphi(z) d\psi(z) \rightarrow \infty \text{ as } y \rightarrow \infty.$$

**Proof of Lemma 2.** We start by the case  $\alpha < \beta$ . Let  $2\delta = \beta - \alpha > 0$ , take  $y \geq A$  so large that  $y^{-\alpha-\delta} \varphi(y) \leq 1$  and write for  $Y > y$ :

$$\int_y^Y \frac{\varphi(z) d\psi(z)}{\varphi(y) \psi(y)} = \int_y^Y \left[ \frac{z^{-\alpha-\delta} \varphi(z)}{y^{-\alpha-\delta} \varphi(y)} - \left( \frac{z}{y} \right)^{-\delta} \right] \frac{z^{\alpha+\delta} d\psi(z)}{y^{\alpha+\delta} \psi(y)} + \int_y^Y \frac{z^\alpha d\psi(z)}{y^\alpha \psi(y)}. \quad (27)$$

Let now

$$S(y) = \sup_{\lambda \geq 1} \left| \frac{(\lambda y)^{-\alpha-\delta} \varphi(\lambda y)}{y^{-\alpha-\delta} \varphi(y)} - \lambda^{-\delta} \right|.$$

A uniform convergence result for the function  $y \mapsto y^{-\alpha-\delta} \varphi(y)$  (see e.g. Theorem 1.5.2 in Bingham et al., 1987) entails that  $S(y)$  converges to 0 as  $y \rightarrow \infty$ . Furthermore, since  $\psi$  is nonincreasing,

$$\left| \int_y^Y \left[ \frac{z^{-\alpha-\delta} \varphi(z)}{y^{-\alpha-\delta} \varphi(y)} - \left( \frac{z}{y} \right)^{-\delta} \right] z^{\alpha+\delta} d\psi(z) \right| \leq -S(y) \int_y^Y z^{\alpha+\delta} d\psi(z). \quad (28)$$

Besides, Theorem 1.6.5 in Bingham et al. (1987) entails for all  $\theta < \beta$

$$\int_y^\infty z^\theta d\psi(z) = -\frac{\beta}{\beta - \theta} y^\theta \psi(y) (1 + o(1)) \text{ as } y \rightarrow \infty. \quad (29)$$

Using (29) with  $\theta = \alpha + \delta$  thus entails that the expression on the left-hand side of (28) has a finite limit as  $Y \rightarrow \infty$  and we may write

$$\left| \int_y^\infty \left[ \frac{z^{-\alpha-\delta} \varphi(z)}{y^{-\alpha-\delta} \varphi(y)} - \left( \frac{z}{y} \right)^{-\delta} \right] z^{\alpha+\delta} d\psi(z) \right| = o(y^{\alpha+\delta} \psi(y)) \text{ as } y \rightarrow \infty. \quad (30)$$

Combining (27), (29) with  $\theta = \alpha$  and (30) concludes the proof of the first statement.

We turn to the case  $\alpha = \beta$ . Pick an arbitrary  $\mu > 1$ ; since  $\psi$  is nonincreasing on  $[A, \infty)$ , we have for  $y \geq A$ :

$$-\int_y^\infty \frac{\varphi(z) d\psi(z)}{\varphi(y) \psi(y)} \geq -\int_y^{\mu y} \frac{\varphi(z) d\psi(z)}{\varphi(y) \psi(y)}. \quad (31)$$

To control the expression on the right-hand side of (31), write

$$\int_y^{\mu y} \frac{\varphi(z)d\psi(z)}{\varphi(y)\psi(y)} = \int_y^{\mu y} \left[ \frac{\varphi(z)}{\varphi(y)} - \left( \frac{z}{y} \right)^\alpha \right] \frac{d\psi(z)}{\psi(y)} + \int_y^{\mu y} \frac{z^\alpha d\psi(z)}{y^\alpha \psi(y)}. \quad (32)$$

Once again, since  $\psi$  is nonincreasing,

$$\left| \int_y^{\mu y} \left[ \frac{\varphi(z)}{\varphi(y)} - \left( \frac{z}{y} \right)^\alpha \right] \frac{d\psi(z)}{\psi(y)} \right| \leq \sup_{1 \leq \lambda \leq \mu} \left| \frac{\varphi(\lambda y)}{\varphi(y)} - \lambda^\alpha \right| \left| 1 - \frac{\psi(\mu y)}{\psi(y)} \right|.$$

A uniform convergence result for the function  $\varphi$  on compact sets (see e.g. Theorem 1.5.2 in Bingham et al. 1987) and the convergence  $\psi(\mu y)/\psi(y) \rightarrow \mu^{-\beta}$  as  $y \rightarrow \infty$  entail

$$\left| \int_y^{\mu y} \left[ \frac{\varphi(z)}{\varphi(y)} - \left( \frac{z}{y} \right)^\alpha \right] \frac{d\psi(z)}{\psi(y)} \right| \rightarrow 0 \text{ as } y \rightarrow \infty. \quad (33)$$

Further, an integration by parts yields

$$- \int_y^{\mu y} \frac{z^\alpha d\psi(z)}{y^\alpha \psi(y)} = \left( 1 - \frac{(\mu y)^\alpha \psi(\mu y)}{y^\alpha \psi(y)} \right) + \alpha \int_y^{\mu y} \left[ \frac{z^\alpha \psi(z)}{y^\alpha \psi(y)} - 1 \right] \frac{dz}{z} + \alpha \log \mu.$$

Since  $y \mapsto y^\alpha \psi(y)$  is slowly varying at infinity, Theorem 1.5.2 in Bingham et al. (1987) gives

$$- \int_y^{\mu y} \frac{z^\alpha d\psi(z)}{y^\alpha \psi(y)} \rightarrow \alpha \log \mu \text{ as } y \rightarrow \infty. \quad (34)$$

Combine then (31), (32), (33) and (34) to get for  $y$  large enough

$$- \int_y^\infty \frac{\varphi(z)d\psi(z)}{\varphi(y)\psi(y)} \geq - \int_y^{\mu y} \frac{\varphi(z)d\psi(z)}{\varphi(y)\psi(y)} = \alpha \log \mu (1 + o(1)) \quad (35)$$

which, since  $\mu > 1$  is arbitrary, proves that the left-hand side of this inequality tends to infinity as  $y \rightarrow \infty$ . We conclude the proof by noting that from (35), for all  $\lambda > 0$ ,

$$\frac{\int_{\lambda y}^\infty \varphi(z)d\psi(z)}{\int_y^\infty \varphi(z)d\psi(z)} - 1 = - \frac{\int_y^{\lambda y} \varphi(z)d\psi(z)}{\int_y^\infty \varphi(z)d\psi(z)} = - \frac{\varphi(y)\psi(y)}{\int_y^\infty \varphi(z)d\psi(z)} \alpha \log(\lambda)(1 + o(1))$$

converges to 0 as  $y \rightarrow \infty$ , which is what we wanted to prove.  $\blacksquare$

If  $\int_{y_0}^\infty dF(z)/\overline{G}(z) < \infty$ , using Lemma 2 with  $\varphi = 1/\overline{G}$  and  $\psi = \overline{F}$  entails

$$\frac{\sqrt{\overline{F}(y)\overline{G}(y)}}{v(y)} \rightarrow \begin{cases} \sqrt{\gamma_G/(\gamma_G - \gamma_F)} & \text{if } \gamma_F < \gamma_G \\ \infty & \text{if } \gamma_F = \gamma_G \end{cases} \text{ as } y \rightarrow \infty, \quad (36)$$

where as in Theorem 1,

$$v(y) := \overline{F}(y) \left( \int_y^\infty \frac{dF(z)}{\overline{G}(z)} \right)^{-1/2}.$$

**Lemma 3** Assume that  $\bar{F} \in \mathcal{RV}_1(-1/\gamma_F)$ ,  $\bar{G} \in \mathcal{RV}_1(-1/\gamma_G)$  where  $\gamma_F, \gamma_G > 0$ . Then as  $y, t \rightarrow \infty$ :

$$\frac{\bar{F}^*(y)}{\bar{F}(y)\bar{G}(y)} \rightarrow \frac{1}{p} \frac{\gamma_G}{\gamma_F + \gamma_G} \text{ and } \frac{\bar{G}^*(t)}{\bar{G}(t)} \rightarrow \frac{1}{p}.$$

**Proof of Lemma 3.** Pick  $y, t > 0$  and recall from (37) that

$$\bar{F}^*(y) = \frac{1}{p} \int_y^\infty \bar{G}(z) dF(z) \text{ and } \bar{G}^*(t) = \frac{1}{p} \int_t^\infty F(z) dG(z). \quad (37)$$

Noting that  $F(z) \rightarrow 1$  as  $z \rightarrow \infty$  and  $dF(z) = -d\bar{F}(z)$ , the result is then a straightforward consequence of Lemma 2 with  $\varphi = \bar{G}$  and  $\psi = \bar{F}$ . ■

**Lemma 4** For all  $y \geq y_0$ ,  $\mathbb{P}(Y_{N,N}^* \leq y) = \left(1 - p\bar{F}^*(y)\right)^n$ . Consequently, if (M) holds,  $y_n \rightarrow \infty$  and  $nv^2(y_n) \rightarrow \infty$  then  $\mathbb{P}(Y_{N,N}^* \leq y_n) \rightarrow 0$ .

**Proof of Lemma 4.** From Lemma 1, given  $N = m$ , the random variables  $Y_1^*, \dots, Y_m^*$  are independent and identically distributed with cdf  $F^*$ . Therefore for all  $y \geq y_0$ , conditioning on  $N$  gives

$$\mathbb{P}(Y_{N,N}^* \leq y) = \sum_{m=0}^n \binom{n}{m} [pF^*(y)]^m (1-p)^{n-m}$$

which yields the first part of the result. Lemma 3 now entails

$$n \log(1 - p\bar{F}^*(y_n)) = -pn\bar{F}^*(y_n)(1 + o(1)) = -\frac{\gamma_G}{\gamma_F + \gamma_G} n\bar{F}(y_n)\bar{G}(y_n)(1 + o(1)).$$

Use then (36) to obtain that  $n\bar{F}(y_n)\bar{G}(y_n) \rightarrow \infty$ ; consequently,  $\mathbb{P}(Y_{N,N}^* \leq y_n) \rightarrow 0$ , which concludes the proof. ■

**Lemma 5** If  $y_n \rightarrow \infty$  then

$$\sup_{\substack{1 \leq i \leq N \\ Y_i^* > y_n}} \frac{C^*(Y_i^*)}{\widehat{C}_N^*(Y_i^*)} = O_{\mathbb{P}}(1).$$

**Proof of Lemma 5.** We closely follow the proof of Lemma 1.2 in Stute [26]. For  $0 < s, t \leq 1$  and  $\lambda \geq 1$ , set

$$\bar{H}_n(s, t) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{U_i \leq s, V_i \leq t\}}$$

where  $(U_1, V_1), \dots, (U_n, V_n)$  is a sample of independent copies of a random pair  $(U, V)$  which is uniformly distributed on the unit square. We shall prove that for every  $0 < a, b \leq 1$  and every  $\mu \in (0, 1)$

$$\mathbb{P} \left( \sup_{\substack{a \leq s \leq 1 \\ b \leq t \leq 1}} -\frac{\bar{H}_n(s, t)}{st} \geq -\mu \right) \leq \exp(-nab h(\mu) + 1) \quad (38)$$

where  $h(z) = 1 + z(\log z - 1)$ . To this end, we follow the steps of Lemma 1.1 in Stute [26], whose proof is itself similar to the proof of inequality (5) p.415 in Shorack and Wellner [25]. For an arbitrary  $t \in [0, 1]$  let  $\mathcal{F}_t$  be the  $\sigma$ -algebra

$$\mathcal{F}_t = \sigma(\{V_i \leq t'\}, 1 \leq i \leq n, t \leq t' \leq 1, U_1, \dots, U_n).$$

Then (see e.g. Proposition 3.6.2 in Shorack and Wellner [25]) for every  $s \in [a, 1]$  the process

$$\left(-\frac{\overline{H}_n(s, t)}{st}, \mathcal{F}_t\right)_{b \leq t \leq 1}$$

is a reverse martingale. Since the image of a martingale by a convex function is a submartingale, it follows that for every  $r > 0$  the process

$$\left(\sup_{a \leq s \leq 1} \exp\left(-r \frac{\overline{H}_n(s, t)}{st}\right), \mathcal{F}_t\right)_{b \leq t \leq 1}$$

is a nonnegative reverse submartingale. Doob's maximal inequality for such processes and the fact that the exponential function is increasing then entail

$$\mathbb{P}\left(\sup_{\substack{a \leq s \leq 1 \\ b \leq t \leq 1}} -\frac{\overline{H}_n(s, t)}{st} \geq -\mu\right) \leq \exp(r\mu) \mathbb{E}\left[\sup_{a \leq s \leq 1} \exp\left(-r \frac{\overline{H}_n(s, b)}{sb}\right)\right]. \quad (39)$$

We shall now control the expectation on the right-hand side by another martingale argument: for an arbitrary  $s \in [0, 1]$  let  $\mathcal{G}_s$  be the  $\sigma$ -algebra

$$\mathcal{G}_s = \sigma(\{U_i \leq s'\}, 1 \leq i \leq n, s \leq s' \leq 1, V_1, \dots, V_n).$$

The same ideas used above show that for every  $q > 1$  and  $r > 0$  the process

$$\left(\exp\left(-\frac{r}{q} \frac{\overline{H}_n(s, b)}{sb}\right), \mathcal{G}_s\right)_{a \leq s \leq 1}$$

is a nonnegative reverse submartingale. Doob's maximal inequality for  $q$ -th moments of a submartingale yields

$$\mathbb{E}\left(\sup_{a \leq s \leq 1} \exp\left(-r \frac{\overline{H}_n(s, b)}{sb}\right)\right) \leq \left[\frac{q}{q-1}\right]^q \mathbb{E}\left(\exp\left(-r \frac{\overline{H}_n(a, b)}{ab}\right)\right).$$

Let  $q \rightarrow \infty$  and recall (39) to get, for every  $r > 0$ ,

$$\mathbb{P}\left(\sup_{\substack{a \leq s \leq 1 \\ b \leq t \leq 1}} -\frac{\overline{H}_n(s, t)}{st} \geq -\mu\right) \leq \exp(r\mu + 1) \mathbb{E}\left(\exp\left(-r \frac{\overline{H}_n(a, b)}{ab}\right)\right).$$

Finally, since the expectation on the right-hand side above can be seen as the moment generating function of a binomial random variable at  $-r/nab$ ,

$$\begin{aligned} \mathbb{E}\left(\exp\left(-r \frac{\overline{H}_n(a, b)}{ab}\right)\right) &= (1 - ab + ab \exp(-r/nab))^n \\ &\leq \exp(-nab(1 - \exp(-r/nab))) \end{aligned}$$

with the classical inequality  $(1 - x/n)^n \leq e^{-x}$  valid for every  $x > 0$ . Putting  $r = -nab \log \mu$  gives (38).

We can now prove our result. Note that  $\widehat{C}_N^*(Y_i^*) \geq 1/N \geq 1/n$  so that

$$\begin{aligned} \left\{ \sup_{\substack{1 \leq i \leq N \\ Y_i^* > y_n}} \frac{C^*(Y_i^*)}{\widehat{C}_N^*(Y_i^*)} \geq \lambda \right\} &\subset \left\{ \sup_{\substack{t \geq y_n \\ t: C^*(t) \geq \lambda/n}} \frac{n}{N} \frac{C^*(t)}{\widehat{C}_N^*(t)} \geq \lambda \right\} \\ &\subset \left\{ \sup_{\substack{t \geq y_n \\ t: C^*(t) \geq \lambda/n}} -\frac{N}{n} \frac{\widehat{C}_N^*(t)}{pC^*(t)} \geq -\frac{1}{\lambda p} \right\}. \end{aligned} \quad (40)$$

Remarking that  $t \mapsto \overline{G}((-t)^-) := \lim_{\substack{s \uparrow -t \\ s < -t}} \overline{G}(s)$  is the cdf of  $-T$ , we find that

$$\frac{N}{n} \widehat{C}_N^*(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{Y_i \leq t \leq T_i\}} = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{Y_i \leq t, -T_i \leq -t\}} \quad (41)$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{U_i \leq F(t), V_i \leq \overline{G}(t^-)\}} \quad (42)$$

where for each  $n$ ,  $(U_1, V_1), \dots, (U_n, V_n)$  is a sample of independent copies of a random pair  $(U, V)$  which is uniformly distributed on the unit square. Finally, recall that  $pC^*(t) = F(t)\overline{G}(t)$  to obtain

$$C^*(t) \geq \lambda/n \Rightarrow \overline{G}(t^-) \geq \overline{G}(t) \geq \lambda p/n. \quad (43)$$

Use now (40) together with (41) and (43) to get

$$\left\{ \sup_{\substack{1 \leq i \leq N \\ Y_i^* > y_n}} \frac{C^*(Y_i^*)}{\widehat{C}_N^*(Y_i^*)} \geq \lambda \right\} \subset \left\{ \sup_{\substack{t: F(y_n) \leq F(t) \leq 1 \\ t: \lambda p/n \leq \overline{G}(t^-) \leq 1}} -\frac{\overline{H}_n(F(t), \overline{G}(t^-))}{F(t)\overline{G}(t^-)} \geq -\frac{1}{\lambda p} \right\}.$$

Noting that  $\log(x)/x \leq e^{-1}$  for all  $x > 1$ , inequality (38) now yields, for every  $\lambda > 0$  such that  $\lambda p > 1$ :

$$\mathbb{P} \left( \sup_{\substack{1 \leq i \leq N \\ Y_i^* > y_n}} \frac{C^*(Y_i^*)}{\widehat{C}_N^*(Y_i^*)} \geq \lambda \right) \leq \exp((- \lambda p + \log(\lambda p) + 1)F(y_n) + 1) = O(\exp(-c\lambda))$$

where  $c$  is a positive constant, of which the result follows.  $\blacksquare$

**Lemma 6** *Let  $(X_n)$  be a sequence of positive real-valued random variables such that for every positive nonrandom sequence  $(\delta_n)$  converging to 0, the random sequence  $(\delta_n X_n)$  converges to 0 in probability. Then  $X_n = O_{\mathbb{P}}(1)$ .*

**Proof of Lemma 6.** Assume that the sequence  $(X_n)$  is not stochastically bounded, that is:

$$\exists \varepsilon > 0, \forall M > 0, \exists m \in \mathbb{N}, \mathbb{P}(|X_m| > M) \geq \varepsilon.$$

This entails that there exists a sequence of integers  $(m_k)$ , which we may choose to be increasing, such that  $\forall k \in \mathbb{N} \setminus \{0\}, \mathbb{P}(|X_{m_k}| > k) \geq \varepsilon$ . Let then  $(\delta_n)$  be the sequence defined as  $\delta_n = 1/k$  if  $m_k \leq n < m_{k+1}$ . The sequence  $(\delta_n)$  converges to 0, and we have  $\forall k \in \mathbb{N} \setminus \{0\}, \mathbb{P}(|\delta_{m_k} X_{m_k}| > 1) \geq \varepsilon$  which shows that  $(\delta_n X_n)$  does not converge in probability to 0: Lemma 6 is proven. ■

For an arbitrary Borel measurable function  $\psi$  such that  $\psi(y) \neq 0$  for  $y$  large enough and such that  $z \mapsto \psi(z)/z$  is integrable in a neighborhood of infinity, we define

$$I(\psi, y) = \frac{1}{\psi(y)} \int_y^\infty \frac{\psi(z)}{z} dz.$$

**Lemma 7** *Let  $\psi \in \mathcal{RV}_2(-\alpha, \Delta)$  with  $\alpha > 0$  and  $|\Delta| \in \mathcal{RV}_1(\rho)$ . Then we have*

$$I(\psi, y) = \frac{1}{\alpha} + \frac{1}{\alpha(\alpha - \rho)} \Delta(y)(1 + o(1)) \text{ as } y \rightarrow \infty.$$

**Proof of Lemma 7.** Start by using the change of variables  $z = yu$  to obtain

$$I(\psi, y) - \frac{1}{\alpha} = \int_1^\infty \varphi_y(u) \frac{du}{u^{\alpha+1}} \text{ with } \varphi_y(u) = \exp \left( \int_y^{yu} \frac{\tilde{\Delta}(v)}{v} dv \right) - 1 \quad (44)$$

where  $\tilde{\Delta}$  is asymptotically equivalent to  $\Delta$ . Since we are only interested in the behavior of  $\tilde{\Delta}$  at infinity, we may assume without loss of generality that  $\tilde{\Delta}$  has constant sign. Using the change of variables  $v = yw$ , a uniform convergence result for the regularly varying function  $|\tilde{\Delta}|$  (see Theorem 1.5.2 in Bingham et al. 1987) yields for all  $u \geq 1$ :

$$\varphi_y(u) = \tilde{\Delta}(y) D_\rho(u)(1 + o(1)) \text{ as } y \rightarrow \infty \text{ with } D_\rho(u) = \int_1^u v^{\rho-1} dv. \quad (45)$$

Furthermore, Theorem 1.5.6 in Bingham et al. (1987) entails that there exists  $y_1 > 0$  such that

$$\forall w \geq 1, \forall y \geq y_1, \left| \frac{\tilde{\Delta}(yw)}{\tilde{\Delta}(y)} \right| \leq 2w^{\rho+\alpha/4}.$$

It is then clear that for  $y \geq y_1$ :

$$\forall u \geq 1, \frac{1}{|\tilde{\Delta}(y)|} \left| \int_y^{yu} \frac{\tilde{\Delta}(v)}{v} dv \right| \leq 2D_{\rho+\alpha/4}(u).$$



Choose now  $y_1$  so large that for every  $y \geq y_1$ ,  $|\tilde{\Delta}(y)| < \alpha/2$ . The inequality  $|e^t - 1| \leq |t|e^{|t|}$ , valid for every  $t \in \mathbb{R}$ , entails

$$\forall u \geq 1, \quad u^{-\alpha-1} \sup_{y \geq y_1} \left| \frac{\varphi_y(u)}{\tilde{\Delta}(y)} \right| \leq 2u^{-1-\alpha/2} D_{\rho+\alpha/4}(u). \quad (46)$$

Because the right-hand side of this inequality defines an integrable function on  $[1, \infty)$ , we may combine (44), (45), (46) and apply the dominated convergence theorem to obtain

$$I(\psi, y) = \frac{1}{\alpha} + \tilde{\Delta}(y) \int_1^\infty u^{-\alpha} D_\rho(u) \frac{du}{u} (1 + o(1)) \text{ as } y \rightarrow \infty.$$

Since  $\tilde{\Delta}$  is asymptotically equivalent to  $\Delta$ , the result is then a straightforward consequence of this equality.  $\blacksquare$

**Lemma 8** *Assume that (C) holds. Let  $\rho_{F^*} = \rho_F \vee \rho_G$  and define for all  $y$ ,  $t > 0$ :*

$$\begin{aligned} \Delta_{F^*}(y) &= \left( \frac{1}{1 - \gamma_{F^*} \rho_{F^*}} - \frac{\gamma_F \rho_F}{1 - \gamma_{F^*} \rho_F} \right) \Delta_F(y) + \frac{\Delta_G(y)}{1 - \gamma_{F^*} \rho_{F^*}} \\ \text{and } \Delta_{G^*}(t) &= \frac{\bar{F}(t)}{\gamma_F + \gamma_G} + \Delta_G(t). \end{aligned}$$

*If  $\rho_F \neq \rho_G$  and  $\rho_G \neq -1/\gamma_F$  then, defining  $\rho_{G^*} = (-1/\gamma_F) \vee \rho_G$ , we have  $|\Delta_{F^*}| \in \mathcal{RV}_1(\rho_{F^*})$ ,  $|\Delta_{G^*}| \in \mathcal{RV}_1(\rho_{G^*})$  and  $\bar{F}^* \in \mathcal{RV}_2(-1/\gamma_{F^*}, \Delta_{F^*})$ ,  $\bar{G}^* \in \mathcal{RV}_2(-1/\gamma_G, \Delta_{G^*})$ .*

**Proof of Lemma 8.** Since  $|\Delta_F| \in \mathcal{RV}_1(\rho_F)$ ,  $|\Delta_G| \in \mathcal{RV}_1(\rho_G)$  and  $\bar{F} \in \mathcal{RV}_1(-1/\gamma_F)$ , it is clear that  $|\Delta_{F^*}| \in \mathcal{RV}_1(\rho_{F^*})$  and  $|\Delta_{G^*}| \in \mathcal{RV}_1(\rho_{G^*})$  under the conditions of the result. To prove that  $\bar{F}^*$  and  $\bar{G}^*$  satisfy the desired second-order condition, we first remark that condition (C) entails that  $F$  and  $G$  are differentiable. Furthermore, we obtain from (37) that

$$y \frac{d}{dy} \log \bar{F}^*(y) = -\frac{yF'(y)\bar{G}(y)}{p\bar{F}^*(y)} \text{ and } t \frac{d}{dt} \log \bar{G}^*(t) = -\frac{tF(t)G'(t)}{p\bar{G}^*(t)}.$$

We start by controlling  $\bar{F}^*(y)$ . Using (C) leads to

$$\frac{p\bar{F}^*(y)}{yF'(y)\bar{G}(y)} = \frac{I(\bar{F}\bar{G}, y) - \gamma_F \tilde{\Delta}_F(y) I(\tilde{\Delta}_F \bar{F}\bar{G}, y)}{1 - \gamma_F \tilde{\Delta}_F(y)}$$

where  $\tilde{\Delta}_F$  is asymptotically equivalent to  $\Delta_F$ . The condition  $\rho_F \neq \rho_G$  entails that the function  $\Delta_F + \Delta_G$  has ultimately constant sign and  $|\Delta_F + \Delta_G|$  is asymptotically equivalent to an ultimately monotonic regularly varying function with index  $\rho_{F^*}$ , and thus  $\bar{F}\bar{G}$  satisfies the second-order condition (C).

Furthermore, note that since  $\tilde{\Delta}_F$  has asymptotically constant sign, we have  $I(\tilde{\Delta}_F \overline{FG}, y) = I(|\tilde{\Delta}_F| \overline{FG}, y)$  when  $y$  is large enough. Applying Lemma 7 to the function  $\overline{FG}$ , Theorem 0.6.(a) in Resnick (2007) to the function  $|\tilde{\Delta}_F| \overline{FG}$  and using a Taylor expansion, we obtain after some algebra:

$$y \frac{d}{dy} \log \overline{F^*}(y) = -\frac{1}{\gamma_{F^*}} + \Delta_{F^*}(y)(1 + o(1)) \text{ as } y \rightarrow \infty.$$

An integration completes the proof of the statement on  $\overline{F^*}$ . We now turn to the statement on  $\overline{G^*}$ . Define  $\Psi : z \mapsto zG'(z)$  and use **(C)** to obtain that  $\Psi \in \mathcal{RV}_1(-1/\gamma_G)$ . Using again **(C)** entails

$$\frac{p\overline{G^*}(t)}{tF(t)G'(t)} = \frac{\overline{F}(t)}{F(t)}(I(\Psi, t) - I(\overline{G}\Psi, t)) + \frac{I(\overline{G}, t) - \gamma_G \tilde{\Delta}_G(t)I(\tilde{\Delta}_G \overline{G}, t)}{1 - \gamma_G \tilde{\Delta}_G(t)}$$

where  $\tilde{\Delta}_G$  is asymptotically equivalent to  $\Delta_G$ . Moreover, we have  $I(\tilde{\Delta}_G \overline{G}, t) = I(|\tilde{\Delta}_G| \overline{G}, t)$  when  $t$  is large enough, because  $\tilde{\Delta}_G$  has asymptotically constant sign. Apply then Theorem 0.6.(a) in Resnick (2007) to the functions  $\Psi$ ,  $\overline{F}\Psi$  and  $|\tilde{\Delta}_G| \overline{G}$ , Lemma 7 to the function  $\overline{G}$  and use a Taylor expansion to get after some computations:

$$t \frac{d}{dt} \log \overline{G^*}(t) = -\frac{1}{\gamma_G} + \Delta_{G^*}(t)(1 + o(1)) \text{ as } t \rightarrow \infty.$$

An integration completes the proof. ■

**Lemma 9** *Let  $\gamma > 0$  and  $Z$  be a random variable whose survival function  $\Psi$  belongs to  $\mathcal{RV}_1(-1/\gamma)$ . Assume that:*

- $N := N(n)$  is a sequence of integer-valued random variables such that there exists a positive sequence  $(u_n)$  of integers tending to infinity with  $N/u_n \xrightarrow{\mathbb{P}} \infty$ ;
- $\hat{\gamma}_N(k_N)$  is a random variable such that the distribution of  $\hat{\gamma}_N(k_N)$  given  $N = m$  is that of

$$\tilde{\gamma}_m(k_m) = \frac{1}{k_m} \sum_{i=1}^{k_m} \log \frac{Z_{m-i+1,m}}{Z_{m-k_m,m}}$$

where  $Z_{1,m} \leq \dots \leq Z_{m,m}$  are the order statistics related to a sample of independent and identically distributed copies  $Z_1, \dots, Z_m$  of  $Z$ .

Then for every sequence  $(k_n)$  such that  $k_n \rightarrow \infty$  and  $k_n/n \rightarrow 0$ , we have  $\hat{\gamma}_N(k_N) \xrightarrow{\mathbb{P}} \gamma$ . Assume further that  $\Psi \in \mathcal{RV}_2(-1/\gamma, \Delta)$ ; then if  $k_n \rightarrow \infty$ ,  $k_n/n \rightarrow 0$  and  $k_n \Delta^2(U(n/k_n)) \rightarrow 0$  where  $U$  is the left-continuous inverse of  $1/\Psi$ , the random variable  $\sqrt{k_N}(\hat{\gamma}_N(k_N) - \gamma)$  is asymptotically Gaussian centered with variance  $\gamma^2$ .

**Proof of Lemma 9.** Note that if  $I_n = [u_n, \infty)$  then  $\mathbb{P}(N \in I_n) \rightarrow 1$ . To prove the consistency statement, we write for all  $t > 0$ ,  $\mathbb{P}(|\hat{\gamma}_N(k_N) - \gamma| > t) = r_{1,n} + r_{2,n}$  with

$$\begin{aligned} r_{1,n} &= \mathbb{P}(|\hat{\gamma}_N(k_N) - \gamma| > t, N \notin I_n) \\ \text{and } r_{2,n} &= \sum_{m \in I_n} \mathbb{P}(|\hat{\gamma}_N(k_N) - \gamma| > t | N = m) \mathbb{P}(N = m). \end{aligned}$$

It is clear that  $r_{1,n} \leq \mathbb{P}(N \notin I_n) \rightarrow 0$ . Besides

$$r_{2,n} \leq \sup_{m \in I_n} \mathbb{P}(|\tilde{\gamma}_m(k_m) - \gamma| > t) \rightarrow 0$$

since  $\tilde{\gamma}_n(k_n) \xrightarrow{\mathbb{P}} \gamma$ , see Theorem 3.2.2 p.70 in de Haan and Ferreira (2006), so that the estimator is indeed consistent. To show the asymptotic normality of  $\hat{\gamma}_N(k_N)$ , we let  $\varphi$  be the cdf of a centered Gaussian distribution with variance  $\gamma^2$  and we write for all  $t \in \mathbb{R}$ ,  $\mathbb{P}(\sqrt{k_N}(\hat{\gamma}_N(k_N) - \gamma) \leq t) - \varphi(t) = r_{3,n} + r_{4,n}$  with

$$\begin{aligned} r_{3,n} &= \mathbb{P}(\sqrt{k_N}(\hat{\gamma}_N(k_N) - \gamma) \leq t, N \notin I_n) - \varphi(t) \mathbb{P}(N \notin I_n) \\ \text{and } r_{4,n} &= \sum_{m \in I_n} \left( \mathbb{P}(\sqrt{k_N}(\hat{\gamma}_N(k_N) - \gamma) \leq t | N = m) - \varphi(t) \right) \mathbb{P}(N = m). \end{aligned}$$

Again, clearly  $|r_{3,n}| \leq 2\mathbb{P}(N \notin I_n) \rightarrow 0$ . Moreover

$$|r_{4,n}| \leq \sup_{m \in I_n} \left| \mathbb{P}(\sqrt{k_m}(\tilde{\gamma}_m(k_m) - \gamma) \leq t) - \varphi(t) \right|$$

so that it is enough to show that  $\tilde{\gamma}_n(k_n)$  converges in distribution to a centered Gaussian distribution with variance  $\gamma^2$ . To this end first note that since  $\Psi$  is continuous, we have for all  $y > 1$  that  $(1/\Psi)(U(y)) = y$  and thus

$$\forall y > 1, \log y = -\log c + \frac{1}{\gamma} \log U(y) - \int_1^{U(y)} \frac{\tilde{\Delta}(v)}{v} dv$$

with  $c > 0$  and  $\tilde{\Delta}$  asymptotically equivalent to  $\Delta$ . Differentiating and carrying out some computations, we get

$$\frac{U'(y)}{U(y)} = \frac{\gamma}{y} + \frac{\bar{\Delta} \circ U(y)}{y} \text{ with } \bar{\Delta}(y) = \frac{\gamma^2 \tilde{\Delta}}{1 - \gamma \tilde{\Delta}}.$$

As a consequence

$$\forall y > 1, U(y) = y^\gamma \left( \bar{c} \exp \left[ \int_1^y \frac{\bar{\delta}(v)}{v} dv \right] \right)$$

where  $\bar{c} > 0$  and  $\bar{\delta} = \bar{\Delta} \circ U$ . Since  $(Z_1, \dots, Z_n)$  has the same distribution as  $(U(X_1), \dots, U(X_n))$  where  $(X_1, \dots, X_n)$  is a sample of independent copies of

standard Pareto random variables, we may then write  $\tilde{\gamma}_n(k_n) \stackrel{d}{=} S_{n,1} + S_{n,2}$  with

$$S_{n,1} = \frac{\gamma}{k_n} \sum_{i=1}^{k_n} \log \frac{X_{n-i+1,n}}{X_{n-k_n,n}} \text{ and } S_{n,2} = \frac{1}{k_n} \sum_{i=1}^{k_n} \int_{X_{n-k_n,n}}^{X_{n-i+1,n}} \frac{\bar{\delta}(v)}{v} dv.$$

Rényi's representation and Paul Lévy's central limit theorem entail

$$\sqrt{k_n}(S_{n,1} - \gamma) \stackrel{d}{=} \gamma \sqrt{k_n} \left( \frac{1}{k_n} \sum_{i=1}^{k_n} \log(X_i) - 1 \right) \xrightarrow{d} \mathcal{N}(0, \gamma^2).$$

Moreover, since  $(n/k_n)X_{n-k_n,n} \xrightarrow{\mathbb{P}} 1$  (see Corollary 2.2.2 p.41 in de Haan and Ferreira 2006) and  $|\bar{\delta}|$  is asymptotically equivalent to the regularly varying and ultimately nonincreasing function  $|\delta| := \gamma^2 |\Delta \circ U|$ , we get with arbitrarily large probability

$$\sqrt{k_n}|S_{n,2}| \leq \sqrt{k_n}|\delta(X_{n-k_n,n})|S_{n,1}(1 + o_{\mathbb{P}}(1)) = O_{\mathbb{P}}(\sqrt{k_n}|\delta(n/k_n)|) \xrightarrow{\mathbb{P}} 0$$

as  $n \rightarrow \infty$ . Using Slutsky's lemma completes the proof of Lemma 9. ■

$p$	0.7			0.8			0.9			0.95		
$\theta$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$\gamma_F = 1/4$	0.08	0.10	0.16	0.07	0.08	0.12	0.06	0.07	0.10	0.05	0.06	0.08
$\gamma_F = 1/2$	0.31	0.38	0.60	0.26	0.31	0.45	0.23	0.27	0.36	0.21	0.25	0.32
$\gamma_F = 1$	1.22	1.53	2.27	1.05	1.28	1.82	0.91	1.08	1.49	0.85	0.99	1.29

$p$	0.7			0.8			0.9			0.95		
$\theta$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$\gamma_F = 1/4$	0.11	0.17	0.36	0.09	0.12	0.25	0.07	0.1	0.16	0.07	0.09	0.13
$\gamma_F = 1/2$	0.34	0.46	0.85	0.28	0.36	0.55	0.24	0.29	0.43	0.22	0.27	0.37
$\gamma_F = 1$	1.22	1.55	2.46	1.05	1.29	1.9	0.92	1.09	1.46	0.86	1.01	1.33

Table 1: Errors associated with the estimator  $\hat{q}_N^{(r(\theta))}$  for  $\delta = 1/3$  (top) and  $\delta = 1$  (bottom).

$p$	0.7			0.8			0.9			0.95		
$\theta$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$\gamma_F = 1/4$	0.004	0.03	0.22	0.003	0.02	0.10	0.002	0.01	0.06	0.002	0.01	0.04
$\gamma_F = 1/2$	0.01	0.10	0.50	0.007	0.05	0.27	0.004	0.03	0.16	0.004	0.03	0.12
$\gamma_F = 1$	0.04	0.39	1.71	0.03	0.25	1.15	0.02	0.13	0.61	0.01	0.09	0.39

$p$	0.7			0.8			0.9			0.95		
$\theta$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$\gamma_F = 1/4$	0.05	0.22	2.84	0.04	0.17	1.00	0.03	0.12	0.49	0.03	0.10	0.30
$\gamma_F = 1/2$	0.04	0.24	2.43	0.03	0.14	0.85	0.02	0.09	0.42	0.02	0.07	0.27
$\gamma_F = 1$	0.05	0.46	2.65	0.03	0.25	1.42	0.02	0.15	0.66	0.02	0.11	0.53

Table 2: Errors associated with the estimator  $\hat{q}_N^{W,(s(\theta))}(.|\alpha_{opt}^{(s(\theta))})$  for  $\delta = 1/3$  (top) and  $\delta = 1$  (bottom).

$p$	0.7			0.8			0.9			0.95		
$\theta$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$\gamma_F = 1/4$	0.10	0.15	0.21	0.07	0.10	0.14	0.05	0.07	0.10	0.05	0.06	0.08
$\gamma_F = 1/2$	0.36	0.51	0.69	0.26	0.36	0.51	0.21	0.27	0.37	0.20	0.24	0.32
$\gamma_F = 1$	1.43	1.97	2.64	1.00	1.42	1.95	0.85	1.06	1.48	0.80	0.98	1.31

$p$	0.7			0.8			0.9			0.95		
$\theta$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$\gamma_F = 1/4$	0.19	0.33	0.53	0.10	0.18	0.31	0.07	0.1	0.18	0.06	0.08	0.13
$\gamma_F = 1/2$	0.46	0.71	1.03	0.29	0.45	0.69	0.23	0.30	0.46	0.21	0.26	0.38
$\gamma_F = 1$	1.43	2.13	2.91	1.06	1.48	2.09	0.86	1.10	1.53	0.81	0.99	1.33

Table 3: Errors associated with the estimator  $\hat{q}_N^{(r^*(\theta)),*}$  for  $\delta = 1/3$  (top) and  $\delta = 1$  (bottom).

$p$	0.7			0.8			0.9			0.95		
$\theta$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$\gamma_F = 1/4$	0.04	0.08	0.14	0.01	0.04	0.08	0.003	0.01	0.04	0.002	0.01	0.03
$\gamma_F = 1/2$	0.12	0.28	0.49	0.04	0.14	0.29	0.007	0.04	0.15	0.004	0.03	0.10
$\gamma_F = 1$	0.5	1.11	1.90	0.12	0.52	1.14	0.03	0.18	0.61	0.01	0.12	0.45

$p$	0.7			0.8			0.9			0.95		
$\theta$	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
$\gamma_F = 1/4$	0.14	0.26	0.44	0.06	0.14	0.25	0.03	0.08	0.16	0.02	0.06	0.14
$\gamma_F = 1/2$	0.22	0.42	0.75	0.07	0.20	0.41	0.02	0.08	0.21	0.02	0.06	0.16
$\gamma_F = 1$	0.47	1.16	2.04	0.16	0.53	1.15	0.03	0.21	0.60	0.02	0.12	0.43

Table 4: Errors associated with the estimator  $\hat{q}_N^{W,(s^*(\theta)),*}(\cdot|\alpha_{opt}^{(s^*(\theta)),*})$  for  $\delta = 1/3$  (top) and  $\delta = 1$  (bottom).

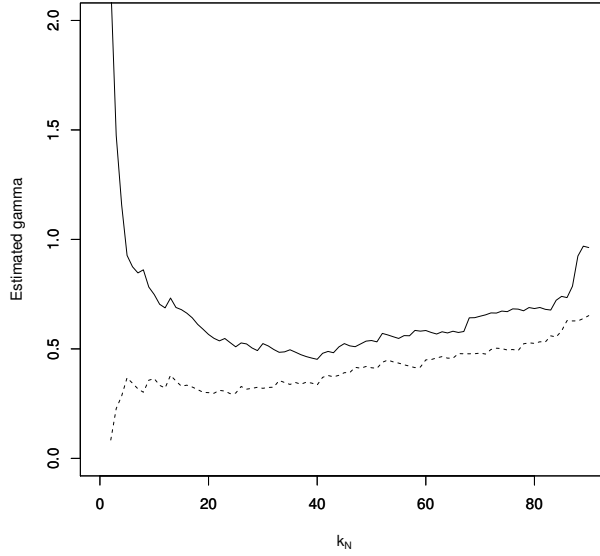


Figure 1: Hill estimators of  $\gamma_G$  (full line) and  $\gamma_{F^*}$  (dashed line) as functions of  $k_N$ .

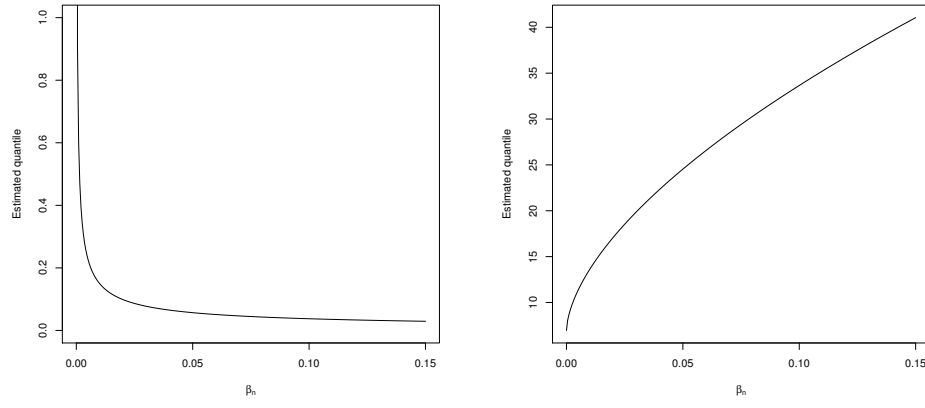


Figure 2: Estimated quantiles for the transformed data (left) and the original data (right).