WORKING PAPERS

N° 1444

November 2023

“Bias-reduced and variance-corrected asymptotic Gaussian Inference about extreme expectiles”

Abdelaati Daouia, Gilles Stupfler and Antoine Usseglio-Carleve

Toulouse School of Economics

Bias-reduced and variance-corrected asymptotic Gaussian inference about extreme expectiles

Abdelaati Daouia[1†], Gilles Stupfler[2*†], Antoine Usseglio-Carleve[3†]

[1]Toulouse School of Economics, University of Toulouse Capitole, France
ORCID: 0000-0003-2621-8860.

[2*]Univ Angers, CNRS, LAREMA, SFR MATHSTIC, F-49000 Angers, France
ORCID: 0000-0003-2497-9412.

[3]Avignon Université, LMA UPR 2151, 84000 Avignon, France
ORCID: 0000-0002-8148-3758.

*Corresponding author(s). E-mail(s): gilles.stupfler@univ-angers.fr;
Contributing authors: abdelaati.daouia@tse-fr.eu;
antoine.usseglio-carleve@univ-avignon.fr;
†These authors contributed equally to this work.

Abstract

The expectile is a prime candidate for being a standard risk measure in actuarial and financial contexts, for its ability to recover information about probabilities and typical behavior of extreme values, as well as its excellent axiomatic properties. A series of recent papers has focused on expectile estimation at extreme levels, with a view on gathering essential information about low-probability, high-impact events that are of most interest to risk managers. The obtention of accurate confidence intervals for extreme expectiles is paramount in any decision process in which they are involved, but actual inference on these tail risk measures is still a difficult question due to their least squares nature and sensitivity to tail heaviness. This article focuses on asymptotic Gaussian inference about tail expectiles in the challenging context of heavy-tailed observations. We use an in-depth analysis of the proofs of asymptotic normality results for two classes of extreme expectile estimators to derive bias-reduced and variance-corrected Gaussian confidence intervals. These, unlike previous attempts in the literature, are well-rooted in statistical theory and can accommodate underlying distributions that display a wide range of tail behaviors. A large-scale simulation study and three real data analyses confirm the versatility of the proposed technique.

Keywords: Asymptotic normality, Bias correction, Expectiles, Extreme values, Inference, Variance correction

1 Introduction

The problem of correctly estimating and inferring extreme risk, carried by low-probability high-impact events such as systemic financial crises and natural disasters, arises in a large range of applications such as insurance and finance, where it is crucial to correctly evaluate and manage the risk carried by a portfolio of claimants or stocks. Several risk measures have been fruitfully used in this context. Among these, expectiles have recently grown increasingly popular for a number of reasons, including the fact that they induce the only law-invariant, coherent (Artzner et al., 1999) and elicitable (Gneiting, 2011) risk measure, see Bellini et al. (2014) and Ziegel (2016). Consequently, a straightforward backtesting methodology can be developed for expectiles, which allows one to rank expectile forecasts by their accuracy (see Theorem 10 in Gneiting, 2011). This crucially hinges on their formulation as minimizers of an asymmetric squared loss function (Newey and Powell, 1987), as follows:

$$\xi_\tau = \underset{\theta \in \mathbb{R}}{\arg\min}\, \mathbb{E}(\eta_\tau(Y-\theta) - \eta_\tau(Y)), \qquad (1.1)$$

where Y denotes the loss variable, assumed to have a finite first moment, $\tau \in (0,1)$ is the asymmetry level, and $\eta_\tau(u) = |\tau - \mathbb{1}\{u \leq 0\}|u^2$ is the so-called expectile check function (throughout, $\mathbb{1}\{\cdot\}$ denotes the indicator function).

Expectile estimation and inference was first developed in Newey and Powell (1987) in the context of testing for homoskedasticity and conditional symmetry in linear regression problems. Recent contributions include deep asymptotic results for the estimation of central, non-tail expectiles of fixed order τ, see for example Holzmann and Klar (2016) and Krätschmer and Zähle (2017). The problem of estimating extreme expectiles, whose order $\tau \uparrow 1$, is more difficult and has been studied only fairly recently, even though it constitutes the right framework for the assessment of extreme risk. The starting point for extreme expectile estimation appears to be a series of papers focusing on the challenging case when the underlying distribution has a heavy right tail (Daouia et al., 2018, 2019, 2020). In this series of articles, two classes of estimators are developed: the first class extrapolates to the far tail a Least Asymmetrically Weighted Squares (LAWS) estimator, obtained through the minimization of the empirical counterpart of problem (1.1), and the second extrapolates a quantile-based (or indirect) estimator whose construction rests upon a remarkable asymptotic proportionality relationship linking extreme expectiles to their quantile analogs.

It was noted later by Padoan and Stupfler (2022) that asymptotic Gaussian inference of extreme expectiles using these two classes of estimators was a difficult question, due to the fact that the asymptotic variances of the Gaussian limiting distributions of the estimators tend to provide a poor representation of the actual uncertainty in finite samples. A solution put forward in Padoan and Stupfler (2022) is to construct corrected confidence intervals that better approximate this uncertainty. However, because their built-in bias correction operates under the restriction that finite-sample bias due to the second order approximation does not dominate, there is no guarantee that they perform well in the wider heavy tail framework. It is now known that comprehensive bias correction is necessary if a reasonable degree of finite-sample accuracy of the extreme expectile estimators is to be ensured (Girard et al., 2022). The variability of the bias-corrected versions introduced in Girard et al. (2022) is, however, similarly hard to handle using straightforward plug-in estimators of the asymptotic variances arising in the Gaussian limiting distributions, and no satisfactory solution for inference is provided therein. More generally, since the precise form of the corrections of Padoan and Stupfler (2022) is motivated only by intensive simulations in certain models, a reverse engineering of their construction is difficult, thus making a potential extension to other types of distributional tails (such as light tails) impossible.

The contribution of the present paper is to develop corrected Gaussian confidence intervals for extreme expectiles of heavy-tailed distributions, with a rigorous theoretical foundation and whose coverage is close to the nominal confidence level even in moderately large samples. Our method essentially consists in carefully identifying, and then correcting, the approximation errors made in the proofs of the asymptotic normality of the extrapolated LAWS and quantile-based estimators. These approximation errors are typically due to either (i) using the asymptotic connection between extreme quantiles and expectiles while ignoring higher-order error terms, (ii) incorrectly neglecting correlations between two estimators when the asymptotic behavior of one of them dominates, or (iii) employing the delta-method for linearization purposes and, in doing so, incurring variance distortions that are not accounted for. We provide successive corrections for each of these types of errors, resulting in refined variance estimators that converge to the asymptotic variances. In order to do so, we start by carefully approximating the finite-sample variability of the standard (*i.e.* without bias reduction) versions of the LAWS and quantile-based estimators. Then, we remark that the bias correction procedures of Girard et al. (2022) introduce further variability, and we design a simple but nonetheless very effective extra correction of the asymptotic variance that is able to push the coverage of the resulting asymptotic Gaussian confidence intervals close to the nominal rate irrespective of the degree of tail heaviness in the underlying distribution.

Even though our corrections are technically involved and rely on a very careful investigation of the probabilistic behavior of the expectile estimators, they are conceptually very simple and result in confidence intervals that are computationally cheap. In particular, we avoid resorting to bootstrap, which is computationally expensive and known to be difficult to calibrate when the underlying distribution is heavy-tailed: Athreya (1987), Knight (1989) and Hall (1990) show that the traditional bootstrap is not consistent for the distribution of a sample mean, and Angus (1993) proves that it is not consistent either for the distribution of a sample minimum or maximum (*i.e.* an extreme value). Being extensions of the mean and extreme values, tail expectiles are similarly difficult to infer using resampling schemes. A potential consistent alternative is subsample bootstrap, but this has been shown to perform poorly both for inference about the mean (Hall and Jing, 1998) and about the tail index of heavy-tailed distributions (Guillou, 2000). Parametric bootstrap in the spirit of Cornea-Madeira and Davidson (2015) is another possibility, but it has not, to the best of our knowledge, been implemented for heavy-tailed distributions. The approach we propose bypasses all these difficulties in using bootstrap by staying within the familiar realm of asymptotic Gaussian inference.

The outline of this paper is the following. Section 2 spells out in detail our statistical framework as well as the two classes of expectile estimators that we will focus on. We then work out bias-reduced and variance-corrected asymptotic Gaussian confidence intervals built on the LAWS and quantile-based estimators in Sections 3 and 4, respectively, and show that they have asymptotically correct coverage. Section 5 examines their finite-sample performance on simulated data and on three samples of real data from insurance and finance. Appendix A describes the estimators of the second-order parameters used in our implementation. Appendix B gives further details as to how our corrections are calculated. Appendix C contains all necessary proofs and Appendix D provides extra finite-sample results about our simulation study and real data analyses. Our methods are implemented in the freely available R package `Expectrem`, which can be downloaded at https://github.com/AntoineUC/Expectrem.

2 Statistical framework and inferential problem

Let $F : y \mapsto \mathbb{P}(Y \le y)$ and $\overline{F} = 1 - F$ denote, respectively, the distribution and survival functions of the loss variable Y, whose large values represent extreme losses. We focus on heavy-tailed distributions that are commonplace in insurance and finance (Embrechts et al., 1997), which amounts to assuming that $\overline{F}$ is regularly varying with negative index in a neighborhood of $+\infty$, that is, there is $\gamma > 0$ such that $\overline{F}(ty)/\overline{F}(t) \to y^{-1/\gamma}$ as $t \to \infty$ for any $y > 0$. Then $\mathbb{E}|Y| < \infty$ provided $\mathbb{E}|\min(Y, 0)| < \infty$ and $\gamma < 1$, so that the expectile ξ_τ is well-defined by (1.1) for any $\tau \in (0, 1)$.

Let the data points $Y_1, \ldots, Y_n$ be independent copies of Y, and $\tau_n, \tau_n' \uparrow 1$ denote high asymmetry levels such that $n(1 - \tau_n) \to \infty$ and $(1 - \tau_n')/(1 - \tau_n) \to 0$ as $n \to \infty$. It is useful to think of τ_n' as the target expectile level satisfying $n(1 - \tau_n') \to c < \infty$, so that the extreme expectile $\xi_{\tau_n'}$ is located in a region where very few or no data points lie, and of τ_n as an intermediate level, *i.e.* “extreme, but not too much”, so that ξ_{τ_n} can be estimated nonparametrically. Then, according to Daouia et al. (2018), an extreme expectile $\xi_{\tau_n'}$ can be estimated by, first, estimating an intermediate expectile ξ_{τ_n} (by, say, $\overline{\xi}_{\tau_n}$) and the tail index γ (by, say, $\overline{\gamma}$), before combining them to obtain an extrapolated estimator of the form

$$\overline{\xi}^{\star}_{\tau_n'} = \left(\frac{1 - \tau_n'}{1 - \tau_n} \right)^{-\overline{\gamma}} \overline{\xi}_{\tau_n}.$$

A first approach to estimate ξ_{τ_n} is by using the empirical counterpart of the minimization problem (1.1), giving rise to the LAWS estimator $\widehat{\xi}_{\tau_n} = \arg\min_{\theta\in\mathbb{R}} \sum_{i=1}^{n} \eta_{\tau_n}(Y_i - \theta)$. An alternative option is to use the asymptotic proportionality relationship $\xi_{\tau_n} \sim (\gamma^{-1} - 1)^{-\gamma} q_{\tau_n}$ (see Bellini and Di Bernardino, 2017), and to estimate q_{τ_n} therein by $\widehat{q}_{\tau_n} = Y_{n-\lfloor n(1-\tau_n)\rfloor,n}$, where $Y_{1,n} \leq Y_{2,n} \leq \cdots \leq Y_{n,n}$ are the order statistics relative to $Y_1, \ldots, Y_n$. Combined with the Hill estimator (Hill, 1975)

$$\widehat{\gamma}^{\mathrm{H}}_{\tau_n} = \frac{1}{\lfloor n(1-\tau_n)\rfloor} \sum_{i=1}^{\lfloor n(1-\tau_n)\rfloor} \log\left(\frac{Y_{n-i+1,n}}{Y_{n-\lfloor n(1-\tau_n)\rfloor,n}}\right),$$

this results in a quantile-based (or indirect) estimator $\widetilde{\xi}_{\tau_n} = (1/\widehat{\gamma}^{\mathrm{H}}_{\tau_n} - 1)^{-\widehat{\gamma}^{\mathrm{H}}_{\tau_n}} \widehat{q}_{\tau_n}$. An expectile-based alternative estimator of γ, pioneered by Girard et al. (2022) and Girard et al. (2022), is

$$\widehat{\gamma}^{\mathrm{E}}_{\tau_n} = \left(1 + \frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}\right)^{-1}$$

where $\widehat{\overline{F}}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{Y_i > x\}$. All in all, the present article will focus on inference about $\xi_{\tau_n'}$ based on the estimators

$$\widehat{\xi}^{\star}_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}^{\mathrm{E}}_{\tau_n}} \widehat{\xi}_{\tau_n} \text{ and} \tag{2.1}$$

$$\widetilde{\xi}^{\star}_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}^{\mathrm{H}}_{\tau_n}} \widetilde{\xi}_{\tau_n} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}^{\mathrm{H}}_{\tau_n}} (1/\widehat{\gamma}^{\mathrm{H}}_{\tau_n} - 1)^{-\widehat{\gamma}^{\mathrm{H}}_{\tau_n}} \widehat{q}_{\tau_n}. \tag{2.2}$$

Their asymptotic analysis requires quantifying the gap between the extremes of Y and pure Pareto extremes through the following second-order regular variation condition.

$\mathcal{C}_2(\gamma, \rho, A)$ There exist $\gamma > 0$, $\rho \leq 0$ and a measurable auxiliary function A having constant sign and converging to 0 at infinity such that for all $y > 0$,

$$\frac{1}{A(1/\overline{F}(t))}\left(\frac{\overline{F}(ty)}{\overline{F}(t)} - y^{-1/\gamma}\right) \to y^{-1/\gamma}\frac{y^{\rho/\gamma}-1}{\gamma\rho}$$

as $t \to \infty$. Here and throughout the ratio $(y^a - 1)/a$ should be read as $\log y$ when $a = 0$.

In this context it was shown in Daouia et al. (2018) and Girard et al. (2022) that, if $\rho < 0$ and under the conditions $\lambda_1 = \lim_{n\to\infty} \sqrt{n(1-\tau_n)} A((1-\tau_n)^{-1}) \in \mathbb{R}$, $\lambda_2 = \lim_{n\to\infty} \sqrt{n(1-\tau_n)}/q_{\tau_n} \in \mathbb{R}$ and $\sqrt{v_n(\tau_n, \tau_n')} := \sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n')) \to \infty$, one has

$$\sqrt{v_n(\tau_n,\tau_n')}\log\left(\frac{\widehat{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}}\right) \xrightarrow{d} \mathcal{N}\left(\frac{\gamma(\gamma^{-1}-1)^{1-\rho}}{1-\gamma-\rho}\lambda_1 + \gamma^2(\gamma^{-1}-1)^{\gamma+1}\mathbb{E}(Y)\lambda_2, \frac{\gamma^3(1-\gamma)}{1-2\gamma}\right) \tag{2.3}$$

when $\mathbb{E}|\min(Y,0)|^2 < \infty$ and $\gamma < 1/2$, and

$$\sqrt{v_n(\tau_n,\tau_n')}\log\left(\frac{\widetilde{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}}\right) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda_1}{1-\rho}, \gamma^2\right). \tag{2.4}$$

The asymptotic distributions in (2.3) and (2.4) feature bias components due to the semiparametric heavy tail framework. This has motivated the recent work of Girard et al. (2022) on bias reduction procedures for extreme expectile estimation. They assume that $A(t) = b\gamma t^{\rho}$, for a certain constant $b \neq 0$ and $\rho < 0$; this condition is satisfied by most classical continuous heavy-tailed distributions, see Table C2. Then, given estimates $\overline{\rho}$ of ρ and $\overline{b}$ of b (we use here the estimators of ρ and b provided by the R package `Expectrem`[1], whose construction we recall in Appendix A), they first introduce a bias-reduced version of the extrapolated LAWS estimator having the form

$$\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}} \widehat{\xi}_{\tau_n} \times (1+\widehat{B}_{1,n})(1+\widehat{B}_{2,n})(1+\widehat{B}_{3,n}), \tag{2.5}$$

where $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$ is a bias-reduced version of $\widehat{\gamma}^{\mathrm{E}}_{\tau_n}$ defined as

$$\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n} = \left(1 + \frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}\left(1 - \frac{\overline{Y}_n}{\widehat{\xi}_{\tau_n}}\right)^{-1}\right.$$

[1] Credit goes to B.G. Manjunath and F. Caeiro for the original R implementation in package `evt0`, which is unfortunately unavailable from CRAN as of April 2023.

$$\times(2\tau_n-1)\left(1+\frac{\bar{b}(\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n}))^{-\bar{\rho}}}{1-\widehat{\gamma}^{\mathrm{E}}_{\tau_n}-\bar{\rho}}\right)\Bigg)^{-1},$$

with $\overline{Y}_n$ being the sample mean of $Y_1,\ldots,Y_n$. [In Girard et al. (2022), a bias-reduced version of the Hill estimator is used in place of $\widehat{\gamma}^{\mathrm{E}}_{\tau_n}$ in the right-hand side above; the R package `Expectrem` has been updated to reflect this change, motivated by the construction of an estimator which relies on expectiles as the main tool.] The quantities $\widehat{B}_{j,n}$, meanwhile, are defined as follows:

$$\widehat{B}_{1,n}=\frac{((1-\tau_n')/(1-\tau_n))^{-\bar{\rho}}-1}{\bar{\rho}}\,\bar{b}\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}(1-\tau_n)^{-\bar{\rho}},$$

$$1+\widehat{B}_{2,n}=(1+\widehat{r}(\tau_n))^{\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}}$$

$$\times\left(1+\frac{\frac{(1/\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}-1)^{-\bar{\rho}}}{(1+\widehat{r}(\tau_n))^{\bar{\rho}}}-1}{\bar{\rho}}\bar{b}\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}(1-\tau_n)^{-\bar{\rho}}\right)^{-1}$$

where $$1+\widehat{r}(\tau_n)=\left(1-\frac{\overline{Y}_n}{\widehat{\xi}_{\tau_n}}\right)\frac{1}{2\tau_n-1}$$

$$\times\left(1+\frac{\bar{b}(\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n}))^{-\bar{\rho}}}{1-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}-\bar{\rho}}\right)^{-1},$$

and $$1+\widehat{B}_{3,n}=(1+\widehat{r}^{\star}(\tau_n'))^{-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}}$$

$$\times\left(1+\frac{\frac{(1/\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}-1)^{-\bar{\rho}}}{(1+\widehat{r}^{\star}(\tau_n'))^{\bar{\rho}}}-1}{\bar{\rho}}\bar{b}\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}(1-\tau_n')^{-\bar{\rho}}\right)$$

where $$1+\widehat{r}^{\star}(\tau_n')=\left(1-\frac{\overline{Y}_n}{\widehat{\xi}^{\star}_{\tau_n'}}\right)\frac{1}{2\tau_n'-1}$$

$$\times\left(1+\frac{\bar{b}(1/\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}-1)^{-\bar{\rho}}}{1-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}-\bar{\rho}}(1-\tau_n')^{-\bar{\rho}}\right)^{-1}.$$

Besides, a bias-reduced version of the indirect, extrapolated quantile-based estimator is given by

$$\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}=\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}^{\mathrm{H,BR}}_{\tau_n}}\{(1/\widehat{\gamma}^{\mathrm{H,BR}}_{\tau_n}-1)^{-\widehat{\gamma}^{\mathrm{H,BR}}_{\tau_n}}\widehat{q}_{\tau_n}\}$$
$$\times(1+\widetilde{B}_{1,n})(1+\widetilde{B}_{3,n}). \quad (2.6)$$

Here $\widetilde{B}_{1,n}$ is defined as $\widehat{B}_{1,n}$, but with

$$\widehat{\gamma}^{\mathrm{H,BR}}_{\tau_n}=\widehat{\gamma}^{\mathrm{H}}_{\tau_n}\left(1-\frac{\bar{b}}{1-\bar{\rho}}(1-\tau_n)^{-\bar{\rho}}\right),$$

which is a bias-reduced version of the Hill estimator proposed by Caeiro et al. (2005), in place of $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$. Similarly for $\widetilde{B}_{3,n}$, where in addition $\widehat{r}^{\star}(\tau_n')$ is replaced by $\widetilde{r}^{\star}(\tau_n')$, in which $\widehat{\xi}^{\star}_{\tau_n'}$ is replaced by $\widetilde{\xi}^{\star}_{\tau_n'}$. Then, $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ and $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ in (2.5) and (2.6) have the same asymptotic behavior as $\widehat{\xi}^{\star}_{\tau_n'}$ and $\widetilde{\xi}^{\star}_{\tau_n'}$ in (2.3) and (2.4), respectively, the only difference being that they are asymptotically unbiased, see Theorem 2 in Girard et al. (2022). A Gaussian $100(1-\alpha)\%$ asymptotic confidence interval for $\xi_{\tau_n'}$ based on $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ is then

$$\widehat{I}^{(1)}_{\tau_n'}(\alpha)=\left[\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}\exp\left(\pm\sqrt{\frac{\widehat{s}^{2,\mathrm{BR}}_n}{v_n(\tau_n,\tau_n')}}z_{1-\alpha/2}\right)\right]$$
$$\text{with } \widehat{s}^{2,\mathrm{BR}}_n=\frac{(\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})^3(1-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})}{1-2\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}}, \quad (2.7)$$

where $z_{1-\alpha/2}$ is the quantile of level $1-\alpha/2$ for the standard Gaussian distribution. An alternative confidence interval based on $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ is

$$\widetilde{I}^{(1)}_{\tau_n'}(\alpha)=\left[\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}\exp\left(\pm\sqrt{\frac{\widetilde{\sigma}^{2,\mathrm{BR}}_n}{v_n(\tau_n,\tau_n')}}z_{1-\alpha/2}\right)\right]$$
$$\text{with } \widetilde{\sigma}^{2,\mathrm{BR}}_n=(\widehat{\gamma}^{\mathrm{H,BR}}_{\tau_n})^2. \quad (2.8)$$

The rationale behind our inference methods is that the asymptotic results recalled thus far ignore the correlations between the intermediate expectile estimator and the tail index estimator, as well as the inherent variability of the intermediate expectile estimator which can be very substantial in the heavy right tail setup. The intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$ then tend to have poor coverage even in the ideal scenario when the underlying distribution is very close to the Pareto distribution, as shown in the middle panels of Figures 1 and 2 below. By comparing their left and right panels, it can also be inferred that the bias correction is reasonably effective at least for sufficiently large τ_n, whereas the estimation of the variances of $\sqrt{v_n(\tau_n,\tau_n')}\log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ and $\sqrt{v_n(\tau_n,\tau_n')}\log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$, by $\widehat{s}^{2,\mathrm{BR}}_n$ and $\widetilde{\sigma}^{2,\mathrm{BR}}_n$ respectively, is a long way off the truth. It is on the accurate estimation of these variances that our contributions below will focus.

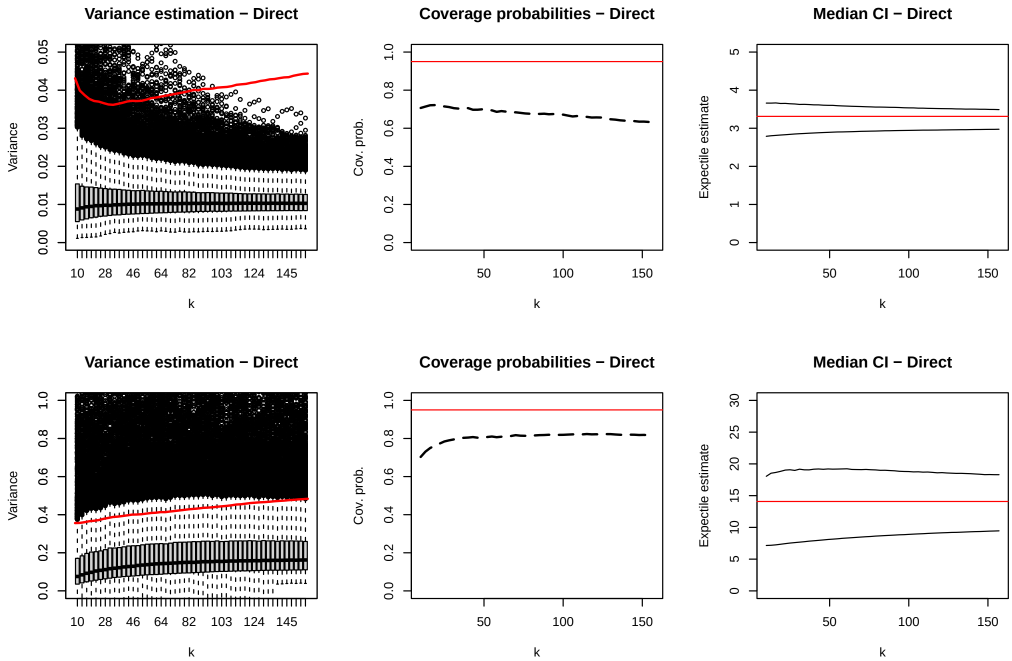

Fig. 1: Left panels: Comparison of the empirical variance of $\sqrt{v_n(\tau_n, \tau_n')}\log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widehat{s}_n^{2,\mathrm{BR}}$. Middle panels: Empirical coverage probability of the associated 95% confidence interval $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$ with the nominal level $1-\alpha = 0.95$ in red. Right panels: Target value $\xi_{\tau_n'}$ (red line) and medians of the lower and upper bounds of $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$. These results were obtained from $N = 5{,}000$ replicated samples of $n = 1{,}000$ independent observations from the following distributions: the Burr distribution with parameters $\gamma = 0.2$ and $\rho = -5$ (first row) and the Fréchet distribution with parameter $\gamma = 0.4$ (second row). The target expectile level is $\tau_n' = 1 - 1/n = 0.999$, and the results are represented as functions of the discretized level $k = k_n = n(1-\tau_n)$ throughout.

3 Bias-reduced and variance-corrected LAWS-based inference

Assume throughout this section that $\mathbb{E}|\min(Y,0)|^2 < \infty$ and $\gamma < 1/2$. The basic idea to show the joint asymptotic normality of $\widehat{\gamma}^{\mathrm{E}}_{\tau_n}$ and $\widehat{\xi}_{\tau_n}$, which are the key building blocks in the estimator $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$, is to note that this amounts to showing the joint asymptotic normality of intermediate expectiles and quantiles at suitably chosen intermediate levels. The crucial point then is that the expectile ξ_τ is just the τth quantile of the distribution function $E = 1 - \overline{E}$ defined as

$$\overline{E}(y) = \frac{\mathbb{E}(|Y-y|\mathbb{1}\{Y > y\})}{\mathbb{E}(|Y-y|)}$$

$$= \frac{\varphi^{(1)}(y)}{2\varphi^{(1)}(y) + y - \mathbb{E}(Y)},$$

with $\varphi^{(a)}(y) = \mathbb{E}([Y-y]^a \mathbb{1}\{Y > y\})$. See Jones (1994). Since $\widehat{\xi}_\tau$ is the τth expectile relative to the empirical distribution function $\widehat{F}_n = 1 - \widehat{\overline{F}}_n$, it is the τth quantile of the distribution function $\widehat{E}_n = 1 - \widehat{\overline{E}}_n$ defined as

$$\widehat{\overline{E}}_n(y) = \frac{\widehat{\varphi}^{(1)}_n(y)}{2\widehat{\varphi}^{(1)}_n(y) + y - \overline{Y}_n}$$

with $\widehat{\varphi}^{(a)}_n(y) = \frac{1}{n}\sum_{i=1}^n (Y_i - y)^a \mathbb{1}\{Y_i > y\}$. It follows that the joint asymptotic normality of intermediate expectiles and quantiles is equivalent to the joint asymptotic normality of $\widehat{\overline{E}}_n$ and $\widehat{\overline{F}}_n$ at suitably chosen intermediate levels. Our purpose

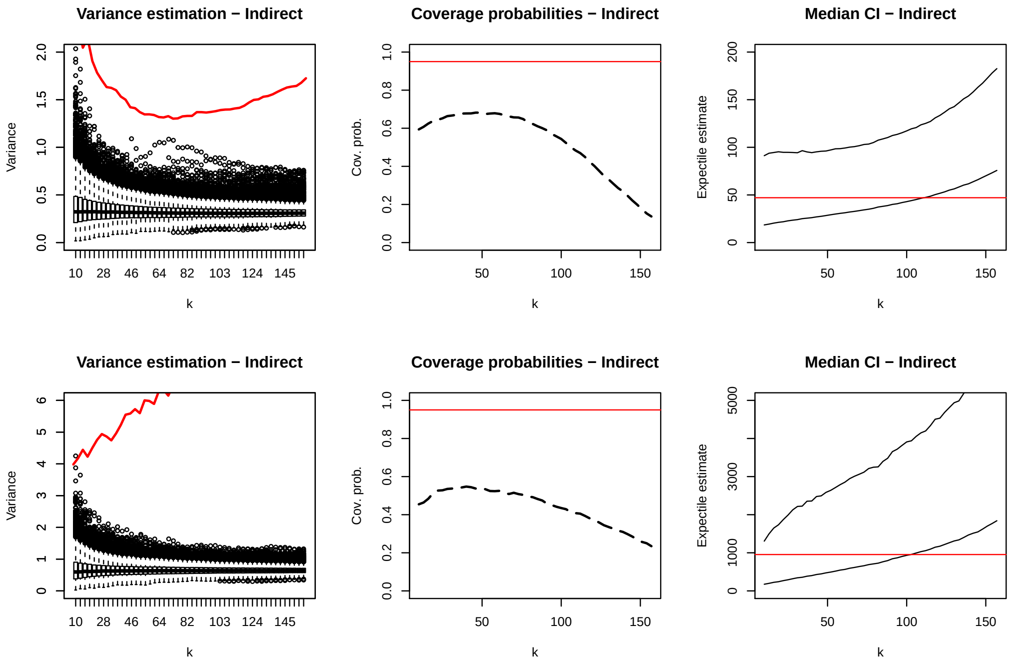

Fig. 2: Left panels: Comparison of the empirical variance of $\sqrt{v_n(\tau_n,\tau_n')}\log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widetilde{\sigma}^{2,\mathrm{BR}}_n$. Middle panels: Empirical coverage probability of the associated 95% confidence interval $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$ with the nominal level $1-\alpha=0.95$ in red. Right panels: Target value $\xi_{\tau_n'}$ (red line) and medians of the lower and upper bounds of $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$. These results were obtained from $N=5{,}000$ replicated samples of $n=1{,}000$ independent observations from the following distributions: the Student distribution with $\nu=1/\gamma=0.6$ degrees of freedom (first row) and the Generalized Pareto Distribution with shape parameter $\gamma=0.8$ and unit scale (second row). The target expectile level is $\tau_n'=1-1/n=0.999$, and the results are represented as functions of the discretized level $k=k_n=n(1-\tau_n)$ throughout.

is to finely quantify the joint uncertainty in this limiting relationship and account for the various sources of errors in the asymptotic approximations, as a preliminary to doing so again for the extrapolated version of $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$. We give a sketch of the most important steps below; full details, following from a careful inspection of the proofs of Theorem 2 in Stupfler and Usseglio-Carleve (2023) and Theorem 2 in Girard et al. (2022), can be found in Appendix B.1.

It is, first of all, straightforward to prove that if τ_n and α_n are intermediate levels (with α_n to be suitably chosen later), the covariance matrix of $\sqrt{n(1-\tau_n)}(\widehat{\varphi}^{(1)}_n(\xi_{\tau_n})/\varphi^{(1)}(\xi_{\tau_n})-1,\widehat{\overline{F}}_n(q_{\alpha_n})/\overline{F}(q_{\alpha_n})-1)$ is *exactly* the 2×2 symmetric matrix $\mathbf{M}^{\varphi}=\mathbf{M}^{\varphi}_n$ having components:

$$M^{\varphi}_{n,11}=(1-\tau_n)\left(\frac{\varphi^{(2)}(\xi_{\tau_n})}{[\varphi^{(1)}(\xi_{\tau_n})]^2}-1\right),$$
$$M^{\varphi}_{n,22}=(1-\tau_n)\frac{\alpha_n}{1-\alpha_n}\ \text{and}\ M^{\varphi}_{n,12}=(1-\tau_n)\times$$
$$\left(\frac{\varphi^{(1)}(\xi_{\tau_n}\vee q_{\alpha_n})+(\xi_{\tau_n}\vee q_{\alpha_n}-\xi_{\tau_n})\overline{F}(\xi_{\tau_n}\vee q_{\alpha_n})}{\varphi^{(1)}(\xi_{\tau_n})(1-\alpha_n)}-1\right).$$

Then, since $\overline{Y}_n-\mathbb{E}(Y)$ converges to 0 at the rate $1/\sqrt{n}$ by the central limit theorem, and thus converges faster than extreme value procedures, we neglect the uncertainty in $\overline{Y}_n$ throughout and in particular in $\widehat{\overline{E}}_n$. A straightforward calculation provides

$$\frac{\widehat{\overline{E}}_n(\xi_{\tau_n})}{\overline{E}(\xi_{\tau_n})} - 1 \approx \frac{(\xi_{\tau_n} - \mathbb{E}(Y))\left(\frac{\widehat{\varphi}_n^{(1)}(\xi_{\tau_n})}{\varphi^{(1)}(\xi_{\tau_n})} - 1\right)}{2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y)}.$$

This suggests that the joint distribution of $\sqrt{n(1-\tau_n)}(\widehat{\overline{E}}_n(\xi_{\tau_n})/\overline{E}(\xi_{\tau_n}) - 1, \widehat{\overline{F}}_n(q_{\alpha_n})/\overline{F}(q_{\alpha_n}) - 1)$ has a covariance matrix that can be more accurately approximated by $\mathbf{M}^E = \mathbf{M}_n^E$ having components:

$$M_{n,11}^E = \frac{(\xi_{\tau_n} - \mathbb{E}(Y))^2 M_{n,11}^{\varphi}}{(2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y))^2},$$
$$M_{n,12}^E = \frac{(\xi_{\tau_n} - \mathbb{E}(Y)) M_{n,12}^{\varphi}}{2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y)}, M_{n,22}^E = M_{n,22}^{\varphi}.$$

The next step is to remark that for $u, v \in \mathbb{R}$,

$$\mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \le u\right\} \cap \left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}} - 1\right) \le v\right\}\right)$$
$$\approx \mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{E}}_n(\xi_{\tau_n})}{\overline{E}(\xi_{\tau_n})} - 1\right) \le \sqrt{n(1-\tau_n)}\left(\frac{\overline{E}(\xi_{\tau_n})}{\overline{E}(y_n)} - 1\right)\right\} \cap \left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(q_{\alpha_n})}{\overline{F}(q_{\alpha_n})} - 1\right) \le \sqrt{n(1-\tau_n)}\left(\frac{\overline{F}(q_{\alpha_n})}{\overline{F}(z_n)} - 1\right)\right\}\right)$$

where $y_n = y_n(u) = \xi_{\tau_n}(1 + u/\sqrt{n(1-\tau_n)})$ and $z_n = z_n(v) = q_{\alpha_n}(1 + v/\sqrt{n(1-\tau_n)})$. A Taylor expansion yields

$$\frac{\overline{E}(\xi_{\tau_n})}{\overline{E}(y_n)} - 1 = \frac{u}{R_n}\frac{1}{\sqrt{n(1-\tau_n)}}(1 + \mathrm{o}(1))$$

with

$$R_n = \frac{\varphi^{(1)}(\xi_{\tau_n})(2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y))}{\xi_{\tau_n}(\varphi^{(1)}(\xi_{\tau_n}) + \overline{F}(\xi_{\tau_n})(\xi_{\tau_n} - \mathbb{E}(Y)))}.$$

Furthermore, assumption $\mathcal{C}_2(\gamma, \rho, A)$ entails $\sqrt{n(1-\tau_n)}(\overline{F}(q_{\alpha_n})/\overline{F}(z_n) - 1) \to v/\gamma$. It readily follows that the covariance matrix of $\sqrt{n(1-\tau_n)}(\widehat{\xi}_{\tau_n}/\xi_{\tau_n} - 1, \widehat{q}_{\alpha_n}/q_{\alpha_n} - 1)$ is approximated by $\mathbf{M}^{\xi} = \mathbf{M}_n^{\xi}$, whose elements are

$$M_{n,11}^{\xi} = R_n^2 M_{n,11}^E, \ M_{n,12}^{\xi} = \gamma R_n M_{n,12}^E \text{ and } M_{n,22}^{\xi} = \gamma^2 M_{n,22}^E.$$

Now that we have obtained a suitable approximation to the covariance matrix of the Gaussian asymptotic distribution of $\widehat{\xi}_{\tau_n}$ and $\widehat{q}_{\alpha_n}$, we recall that $\widehat{\gamma}_{\tau_n}^{\mathrm{E}} = (1 + \widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}$ and we note that

$$\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} - \frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right) \le z\right\}$$
$$= \left\{\widehat{\xi}_{\tau_n} \ge \widehat{q}_{1-(1-\tau_n)(\overline{F}(\xi_{\tau_n})/(1-\tau_n) + z/\sqrt{n(1-\tau_n)})}\right\}.$$

The intermediate quantile level $\beta_n = \beta_n(z) = 1 - (1-\tau_n)(\overline{F}(\xi_{\tau_n})/(1-\tau_n) + z/\sqrt{n(1-\tau_n)})$ in the right-hand side is such that $1 - \beta_n \sim 1 - \alpha_n = \overline{F}(\xi_{\tau_n}) \sim (\gamma^{-1} - 1)(1-\tau_n)$. Then

$$\mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} - \frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right) \le z\right\} \cap \left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \le z'\right\}\right)$$
$$\approx \mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}} - 1\right)\frac{q_{\alpha_n}}{\xi_{\tau_n}} - \sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \le -\sqrt{n(1-\tau_n)}\left(\frac{q_{\beta_n}}{\xi_{\tau_n}} - 1\right)\right\} \cap \left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \le z'\right\}\right).$$

From Proposition 1 in Daouia et al. (2020) and the second-order regular variation assumption,

$$\sqrt{n(1-\tau_n)}\left(\frac{q_{\beta_n}}{\xi_{\tau_n}} - 1\right) = \sqrt{n(1-\tau_n)}\left(\frac{q_{\beta_n}}{q_{\tau_n}} \times \frac{q_{\tau_n}}{\xi_{\tau_n}} - 1\right) \approx -\gamma\frac{1-\tau_n}{\overline{F}(\xi_{\tau_n})}z$$

up to bias terms, which will be taken care of in the bias reduction procedure (and therefore can be neglected at this stage). Conclude that the asymptotic distribution of

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}-\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n},\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)$$

is essentially that of

$$\begin{pmatrix} -\dfrac{1}{\gamma}\dfrac{\overline{F}(\xi_{\tau_n})}{1-\tau_n} & \dfrac{1}{\gamma}\dfrac{\overline{F}(\xi_{\tau_n})}{1-\tau_n} \\ 1 & 0 \end{pmatrix}\sqrt{n(1-\tau_n)}\begin{pmatrix} \dfrac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1 \\ \dfrac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}}-1 \end{pmatrix},$$

and so can be approximated by a centered Gaussian distribution having covariance matrix $\mathbf{M}=\mathbf{M}_n$ defined by its elements

$$M_{n,11}=\frac{1}{\gamma^2}\left(\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right)^2\left[M^{\xi}_{n,11}-2M^{\xi}_{n,12}+M^{\xi}_{n,22}\right],$$
$$M_{n,12}=\frac{1}{\gamma}\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\left[M^{\xi}_{n,12}-M^{\xi}_{n,11}\right]$$

and $M_{n,22}=M^{\xi}_{n,11}$.

It remains to find a precise version of the asymptotic behavior of $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$, and for this we start by considering $\widehat{\gamma}^{\mathrm{E}}_{\tau_n}$. For this purpose, we set $\psi(x)=1/(1+x)$ and note the following refinement of the delta-method as a formal power series expansion of ψ:

$$\begin{aligned}&\sqrt{n(1-\tau_n)}(\widehat{\gamma}^{\mathrm{E}}_{\tau_n}-\psi(\overline{F}(\xi_{\tau_n})/(1-\tau_n)))\\ &=\sqrt{n(1-\tau_n)}(\psi(\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))\\ &\qquad-\psi(\overline{F}(\xi_{\tau_n})/(1-\tau_n)))\\ &=\sum_{j=1}^{\infty}\frac{\psi^{(j)}(\overline{F}(\xi_{\tau_n})/(1-\tau_n))}{j!}[n(1-\tau_n)]^{(1-j)/2}\\ &\qquad\times\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}-\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right)\right\}^j.\end{aligned}$$

Evaluating the variance of the right-hand side and its covariance with $\sqrt{n(1-\tau_n)}(\widehat{\xi}_{\tau_n}/\xi_{\tau_n}-1)$ as formal power series and truncating the resulting series at an order $1/[n(1-\tau_n)]^J$, for a suitably chosen $J\geq 1$, will lead to an accurate approximation $\mathfrak{M}(J)=\mathfrak{M}_n(J)$ of the covariance matrix of $\sqrt{n(1-\tau_n)}(\widehat{\gamma}^{\mathrm{E}}_{\tau_n}-\gamma,\widehat{\xi}_{\tau_n}/\xi_{\tau_n}-1)$. Of course, first of all, $\mathfrak{M}_{n,22}(J)=M_{n,22}$ since this element is the asymptotic variance of $\sqrt{n(1-\tau_n)}(\widehat{\xi}_{\tau_n}/\xi_{\tau_n}-1)$ and does not involve the above power series. Then straightforward calculations based on higher order moments and covariances of Gaussian random variables entail

$$\mathfrak{M}_{n,11}(\infty)=\sum_{j=0}^{\infty}\frac{\mathfrak{m}_{j,11}}{[n(1-\tau_n)]^j}\ \text{with}$$
$$\mathfrak{m}_{j,11}=(1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-2(j+2)}M^{j+1}_{n,11}$$
$$\times\left((2j+1)!!(2j+1)-\sum_{i=1}^{j}(2i-1)!!(2j+1-2i)!!\right)$$

where the double factorial $N!!$ denotes the product of all integers from 1 to N having the same parity as N, and

$$\mathfrak{M}_{n,12}(\infty)=\sum_{j=0}^{\infty}\frac{\mathfrak{m}_{j,12}}{[n(1-\tau_n)]^j}\ \text{with}$$
$$\mathfrak{m}_{j,12}=-(1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-2(j+1)}M^{j}_{n,11}\times M_{n,12}(2j+1)!!.$$

The matrix $\mathfrak{M}(J)=\mathfrak{M}_n(J)$, approximating the covariance matrix of $\sqrt{n(1-\tau_n)}(\widehat{\gamma}^{\mathrm{E}}_{\tau_n}-\psi(\overline{F}(\xi_{\tau_n})/(1-\tau_n)),\widehat{\xi}_{\tau_n}/\xi_{\tau_n}-1)$, is finally obtained by truncating each of the series defining $\mathfrak{M}_{n,11}(\infty)$ and $\mathfrak{M}_{n,12}(\infty)$ at order $1/[n(1-\tau_n)]^J$. In practice $J=1$ already provides reasonable results, and we adopt this choice in our implementation. We use this in order to approximate the uncertainty about $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$. Straightforward calculations yield

$$\begin{aligned}\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}-\gamma&=u_n(\widehat{\gamma}^{\mathrm{E}}_{\tau_n},\widehat{\xi}_{\tau_n}/\xi_{\tau_n})\\ &\quad-u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1},1)\\ &\quad+\mathrm{O}_{\mathbb{P}}(A((1-\tau_n)^{-1}))\end{aligned}$$

where

$$u_n(x,y)=\left(1+(2\tau_n-1)\left(\frac{1}{x}-1\right)\left(1-\frac{\mathbb{E}(Y)}{y\xi_{\tau_n}}\right)^{-1}\right)^{-1}.$$

Here and throughout the paper, we have neglected the finite-sample variability in $\overline{Y}_n$, and we neglect any term proportional to (or dominated by) $A((1-\tau_n)^{-1})$ for the purpose of approximating the variance of our estimators only. The rationale behind this choice is that, since $A(t) = b\gamma t^\rho$, keeping these terms in this kind of calculation would, because of the bias reduction procedure, entail ultimately having to approximate the correlation of estimators of the second-order parameters ρ and b with estimators of other extreme value parameters, here $\widehat{\gamma}^{\mathrm{E}}_{\tau_n}$ and $\widehat{\xi}_{\tau_n}$. This is a joint convergence problem which, to the best of our knowledge, remains open, and whose solution deserves a separate in-depth study. Moreover, in many usual cases in extreme value theory (for instance, when $\gamma < 1/2$), one typically has $\gamma < -\rho$, and then $A(t)$ is negligible with respect to $1/q_{1-1/t} \sim (\gamma^{-1}-1)^{-\gamma}/\xi_{1-1/t}$ as $t \to \infty$, which further justifies this modeling choice in our setting. The random vector

$$\sqrt{n(1-\tau_n)}\left(\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n} - \gamma, \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right)$$

can thus be considered asymptotically Gaussian centered with a covariance matrix approximated by $\mathfrak{M}^{\mathrm{BR}}(J) = \mathfrak{M}^{\mathrm{BR}}_n(J)$ whose elements are

$$\begin{aligned}
&\mathfrak{M}^{\mathrm{BR}}_{n,11}(J)\\
&= (\partial_1 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1},1))^2\mathfrak{M}_{n,11}(J)\\
&+ (\partial_2 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1},1))^2\mathfrak{M}_{n,22}(J)\\
&+ 2\partial_1 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1},1)\\
&\times \partial_2 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1},1)\mathfrak{M}_{n,12}(J),\\
&\mathfrak{M}^{\mathrm{BR}}_{n,12}(J)\\
&= \partial_1 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1},1)\mathfrak{M}_{n,12}(J)\\
&+ \partial_2 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1},1)\mathfrak{M}_{n,22}(J),\\
&\mathfrak{M}^{\mathrm{BR}}_{n,22}(J) = \mathfrak{M}_{n,22}(J).
\end{aligned}$$

Our final step is to combine all these elements in order to accurately quantify the uncertainty in $\log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$. Starting with (2.5), and neglecting any term proportional to $A((1-\tau_n)^{-1})$ or $A((1-\tau_n')^{-1})$ and approximating $\overline{Y}_n$ by $\mathbb{E}(Y)$, we find

$$\begin{aligned}
\log\frac{\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\xi_{\tau_n'}} &\approx \log\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}\\
&+ \left[\log\left(\frac{1-\tau_n}{1-\tau_n'}\right) + \log\left(\frac{2\tau_n'-1}{2\tau_n-1}\right)\right](\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n} - \gamma)\\
&+ \widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}\log\left(1-\frac{\mathbb{E}(Y)}{\widehat{\xi}_{\tau_n}}\right) - \gamma\log\left(1-\frac{\mathbb{E}(Y)}{\xi_{\tau_n}}\right)\\
&- \left[\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}\log\left(1-\frac{\mathbb{E}(Y)}{\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}}\widehat{\xi}_{\tau_n}}\right)\right.\\
&\left.\qquad -\gamma\log\left(1-\frac{\mathbb{E}(Y)}{\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma}\xi_{\tau_n}}\right)\right].
\end{aligned}$$

This eventually leads to the approximation

$$\begin{aligned}
&\sqrt{v_n(\tau_n,\tau_n')}\log\frac{\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\xi_{\tau_n'}}\\
&\approx \sqrt{n(1-\tau_n)}\left\{g_n(\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}, \widehat{\xi}_{\tau_n}/\xi_{\tau_n}) - g_n(\gamma,1)\right\}
\end{aligned}$$

where

$$\begin{aligned}
g_n(x,y) &= x\left(1+\frac{\log((2\tau_n'-1)/(2\tau_n-1))}{\log((1-\tau_n)/(1-\tau_n'))}\right)\\
&+ \frac{\log(y)}{\log((1-\tau_n)/(1-\tau_n'))}\\
&+ \frac{x}{\log((1-\tau_n)/(1-\tau_n'))}\log\left(1-\frac{\mathbb{E}(Y)}{\xi_{\tau_n}y}\right)\\
&- \frac{x}{\log((1-\tau_n)/(1-\tau_n'))}\log\left(1-\frac{\mathbb{E}(Y)}{\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-x}\xi_{\tau_n}y}\right).
\end{aligned}$$

The variance of the asymptotic Gaussian distribution of $\sqrt{v_n(\tau_n,\tau_n')}\log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ is then well approximated by

$$\begin{aligned}
s_n^2(J) &= (\partial_1 g_n(\gamma,1))^2\mathfrak{M}^{\mathrm{BR}}_{n,11}(J)\\
&+ 2\partial_1 g_n(\gamma,1)\partial_2 g_n(\gamma,1)\mathfrak{M}^{\mathrm{BR}}_{n,12}(J)\\
&+ (\partial_2 g_n(\gamma,1))^2\mathfrak{M}^{\mathrm{BR}}_{n,22}(J).
\end{aligned}$$

In order to produce confidence intervals, we estimate γ by $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$, ξ_{τ_n} by $\widehat{\xi}_{\tau_n}$, $\mathbb{E}(Y)$ by $\overline{Y}_n$, $\overline{F}(\xi_{\tau_n})$ by $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$, $\varphi^{(1)}(\xi_{\tau_n})$ by

$$\widehat{\varphi}^{(1)}_n(\widehat{\xi}_{\tau_n}) = \frac{1}{n}\sum_{i=1}^n (Y_i - \widehat{\xi}_{\tau_n})\mathbb{1}\{Y_i > \widehat{\xi}_{\tau_n}\}.$$

The estimation of $\varphi^{(2)}(\xi_{\tau_n})$ is more complex, because its naive empirical counterpart $\widehat{\varphi}^{(2)}_n(\widehat{\xi}_{\tau_n})$ is unbiased but highly skewed, and therefore tends to vastly underestimate $\varphi^{(2)}(\xi_{\tau_n})$. A second-order approximation of the underlying distribution function $\overline{F}$ in a neighborhood of infinity suggests the more accurate estimator

$$\widetilde{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n}) = 2\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})\widehat{\xi}_{\tau_n}^2(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})^2 \times \left(\frac{1}{(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})(1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})} + \frac{\overline{b}(\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n}))^{-\overline{\rho}}}{\overline{\rho}} \times \left\{ \frac{1}{(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}-\overline{\rho})(1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}-\overline{\rho})} - \frac{1}{(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})(1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})} \right\} \right).$$

We deduce from these calculations a corrected asymptotic Gaussian confidence interval for $\xi_{\tau_n'}$ at level $1-\alpha$ as

$$\widehat{I}_{\tau_n'}^{(2)}(\alpha) = \widehat{I}_{\tau_n'}^{(2)}(\alpha; J) = \left[\widehat{\xi}_{\tau_n'}^{\star,\mathrm{BR}} \exp\left(\pm \sqrt{\frac{\widehat{s}_n^2(J)}{v_n(\tau_n,\tau_n')}} z_{1-\alpha/2} \right) \right]$$

where $\widehat{s}_n^2(J)$ is obtained from $s_n^2(J)$ by replacing γ by $\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}$, ξ_{τ_n} by $\widehat{\xi}_{\tau_n}$, $\mathbb{E}(Y)$ by $\overline{Y}_n$, $\overline{F}(\xi_{\tau_n})$ by $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$, $\varphi^{(1)}(\xi_{\tau_n})$ by $\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n})$ and $\varphi^{(2)}(\xi_{\tau_n})$ by $\widetilde{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n})$. The R function `CIextExpect` (with `method="direct"`), available as part of our package `Expectrem`, computes this confidence interval.[2]

We finally state and prove that this corrected confidence interval has asymptotically correct coverage and is always (asymptotically) longer than the naive interval $\widehat{I}_{\tau_n'}^{(1)}(\alpha)$.

Theorem 1. *Assume that* $\mathbb{E}(|\min(Y,0)|^{2+\delta}) < \infty$ *for some* $\delta > 0$*, and condition* $\mathcal{C}_2(\gamma,\rho,A)$ *holds with* $0 < \gamma < 1/2$*,* $\rho < 0$ *and* $A(t) = b\gamma t^{\rho}$*. Let* τ_n*,* τ_n' *be two sequences such that* $\tau_n, \tau_n' \uparrow 1$ *as* $n \to \infty$*,* $n(1-\tau_n) \to \infty$*,* $(1-\tau_n')/(1-\tau_n) \to 0$ *and* $\log((1-\tau_n)/(1-\tau_n'))/\sqrt{n(1-\tau_n)} \to 0$ *as* $n \to \infty$*. Assume further that* $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to 0$ *and* $\sqrt{n(1-\tau_n)}/q_{\tau_n} \to \lambda \in \mathbb{R}$*, and* $\overline{\rho}$ *and* $\overline{b}$ *are consistent estimators of* ρ *and* b *such that* $(\overline{\rho}-\rho)\log(n) = \mathrm{o}_{\mathbb{P}}(1)$*. Then:*

[2] In small samples we compute a very slightly different version based on a modification $\breve{\varphi}_n^{(2)}$ of $\widetilde{\varphi}_n$ for added stability; see Section 5.1 for all necessary details.

(i) For any $J \geq 1$*,* $\widehat{s}_n^2(J) \to \gamma^3(1-\gamma)/(1-2\gamma)$ *in probability, and*

$$\forall \alpha \in (0,1), \quad \lim_{n\to\infty} \mathbb{P}\left(\xi_{\tau_n'} \in \widehat{I}_{\tau_n'}^{(2)}(\alpha; J)\right) = 1-\alpha.$$

(ii) One has

$$\frac{\xi_{\tau_n'}}{\sqrt{n(1-\tau_n)}}(\mathrm{length}(\widehat{I}_{\tau_n'}^{(2)}(\alpha;J)) - \mathrm{length}(\widehat{I}_{\tau_n'}^{(1)}(\alpha))) \to \sqrt{\frac{\gamma^3}{(1-\gamma)(1-2\gamma)}} z_{1-\alpha/2}$$

in probability.

4 Bias and variance-corrected quantile-based inference

For the purpose of carrying out inference using the quantile-based extrapolated estimator $\widetilde{\xi}_{\tau_n'}^{\star,\mathrm{RB}}$, the key is to obtain an accurate representation of the uncertainty in the pair $(\widehat{\gamma}_{\tau_n}^{\mathrm{H}}, \widetilde{\xi}_{\tau_n})$, where the quantile-based intermediate expectile estimator is $\widetilde{\xi}_{\tau_n} = (1/\widehat{\gamma}_{\tau_n}^{\mathrm{H}} - 1)^{-\widehat{\gamma}_{\tau_n}^{\mathrm{H}}} \widehat{q}_{\tau_n}$. Throughout this section we take $\tau_n = 1 - k_n/n$, where k_n is a sequence of integers, $\widehat{q}_{\tau_n} = Y_{n-k_n,n}$ is the corresponding intermediate order statistic, and $\widehat{\gamma}_{\tau_n}^{\mathrm{H}} = \widehat{\gamma}_{1-k_n/n}^{\mathrm{H}}$ is the usual Hill estimator of γ calculated upon the top k_n log-spacings. Again, we only give the main steps of our construction, with full technical details reported in Appendix B.2.

The idea is to write

$$\sqrt{k_n}\log\left(\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n}}\right) = \sqrt{k_n}\{\phi(\widehat{\gamma}_{1-k_n/n}^{\mathrm{H}}) - \phi(\gamma)\} + \sqrt{k_n}\log\left(\frac{\widehat{q}_{1-k_n/n}}{q_{1-k_n/n}}\right)$$

where $\phi(x) = -x\log(x^{-1}-1)$ and, analogously to Section 3, to construct a formal power series expansion of ϕ in a neighborhood of γ:

$$\sqrt{k_n}\log\left(\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n}}\right) = \sqrt{k_n}\log\left(\frac{\widehat{q}_{1-k_n/n}}{q_{1-k_n/n}}\right)$$

$$+\sum_{j=1}^{\infty}\frac{\phi^{(j)}(\gamma)}{j!}k_n^{\frac{1-j}{2}}\left\{\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma)\right\}^j.$$

We approximate the variance of $\sqrt{k_n}\log(\widetilde{\xi}_{1-k_n/n}/((\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n}))$ and its covariance with $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma)$ through a truncated formal power series at an order $1/k_n^J$, for a suitably chosen $J\geq 1$. This provides a finite-sample correction for the asymptotic covariance matrix of $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma,\log(\widetilde{\xi}_{1-k_n/n}/((\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n})))$, and hence a correction for the asymptotic variance of the extrapolated quantile-based extreme expectile estimator. The resulting matrix $\mathbf{V}(J)=\mathbf{V}_n(J)$ has the following coefficients: the first variance term $V_{11}(J)=V_{11}(0)=\gamma^2$, obtained from the asymptotic variance of $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma)$, does not depend on J. The second variance term $V_{22}(J)$, found by evaluating the variance of the right-hand side in the above expansion, is obtained by truncating the series

$$V_{22}(\infty)=\gamma^2+\mathrm{Var}\left(\sum_{j=1}^{\infty}\frac{\phi^{(j)}(\gamma)}{j!}k_n^{\frac{1-j}{2}}\gamma^jZ^j\right)$$

where Z is standard Gaussian. [This is justified because, by Theorem 2.4.8 p.52, Lemma 3.2.3 p.71 and Theorem 3.2.5 p.74 in de Haan and Ferreira (2006), $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma)$ and $\sqrt{k_n}(\widehat{q}_{1-k_n/n}/q_{1-k_n/n}-1)$ are asymptotically Gaussian and independent with asymptotic variance γ^2.] After straightforward calculations, gathering powers of k_n together, we find

$$V_{22}(\infty)=\gamma^2\Bigg(1+(m(\gamma))^2+\sum_{j=1}^{\infty}\frac{\gamma^{2j}}{k_n^j}\left((2j+1)!!\sum_{i=1}^{2j+1}\frac{\phi^{(i)}(\gamma)\phi^{(2j+2-i)}(\gamma)}{i!(2j+2-i)!}-\frac{1}{2^{j+1}}\sum_{i=1}^{j}\frac{\phi^{(2i)}(\gamma)\phi^{(2j+2-2i)}(\gamma)}{i!(j+1-i)!}\right)\Bigg)=:\sum_{j=0}^{\infty}\frac{v_{j,22}}{k_n^j}$$

where $m(\gamma)=\phi'(\gamma)=(1-\gamma)^{-1}-\log(\gamma^{-1}-1)$. An expression of the formal covariance term $V_{12}(\infty)$ is similarly easily derived: if Z is standard Gaussian,

$$V_{12}(\infty)=\sum_{j=0}^{\infty}\frac{\phi^{(j)}(\gamma)}{j!}k_n^{\frac{1-j}{2}}\gamma^{j+1}\mathbb{E}(Z^{j+1})=\sum_{j=0}^{\infty}\frac{\phi^{(2j+1)}(\gamma)}{2^jj!k_n^j}\gamma^{2j+2}=:\sum_{j=0}^{\infty}\frac{v_{j,12}}{k_n^j}.$$

In practice we truncate these power series at order $1/k_n^J$, for a given $J\geq 1$, so that a refined approximation to the covariance matrix of $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma,\log(\widetilde{\xi}_{1-k_n/n}/((\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n})))$ is the symmetric matrix $\mathbf{V}(J)$ whose elements are $V_{11}(J)$, $V_{12}(J)$ and $V_{22}(J)$. In our implementation we stop the approximation at order $1/k_n^2$, that is, $J=2$. Simple expressions of the coefficients $v_{j,12}$ and $v_{j,22}$, for $j=0,1,2$, are provided in Appendix B.2.

We may now proceed with the construction of a confidence interval for $\xi_{\tau_n'}$. We again neglect any term proportional to $A((1-\tau_n)^{-1})$ or $A((1-\tau_n')^{-1})$, and we neglect the finite-sample variability in $\overline{Y}_n$ as an estimator of $\mathbb{E}(Y)$. Starting from Equation (2.6), straightforward calculations lead to the approximation

$$\log\frac{\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\xi_{\tau_n'}}\approx\log\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n}}+\left(\log\left(\frac{k_n}{n(1-\tau_n')}\right)+\log(2\tau_n'-1)\right)(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma)-\left(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}\log\left(1-\frac{\mathbb{E}(Y)}{\widetilde{\xi}^{\star}_{\tau_n'}}\right)-\gamma\log\left(1-\frac{\mathbb{E}(Y)}{\left(\frac{n(1-\tau_n')}{k_n}\right)^{-\gamma}(\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n}}\right)\right).$$

This can be reformulated as

$$\sqrt{v_n(1-k_n/n,\tau_n')}\log\frac{\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\xi_{\tau_n'}}\approx\sqrt{k_n}\left\{h_n\left(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n},\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n}}\right)-h_n(\gamma,1)\right\}$$

where

$$h_n(x,y)=\left(1+\frac{\log(2\tau_n'-1)}{\log(k_n/(n(1-\tau_n')))}\right)x$$

$$+ \frac{\log(y)}{\log(k_n/(n(1-\tau_n')))} - \frac{x}{\log(k_n/(n(1-\tau_n')))}$$
$$\times \log\left(1 - \frac{\mathbb{E}(Y)}{\left(\frac{n(1-\tau_n')}{k_n}\right)^{-x}(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}\, y}\right).$$

This suggests that the variance of $\sqrt{v_n(1-k_n/n,\tau_n')}\log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ is well approximated by

$$\sigma_n^2(J) = (\partial_1 h_n(\gamma,1))^2 V_{11}(J) + 2\partial_1 h_n(\gamma,1)\partial_2 h_n(\gamma,1) V_{12}(J) + (\partial_2 h_n(\gamma,1))^2 V_{22}(J).$$

We deduce from these calculations a corrected asymptotic Gaussian confidence interval for $\xi_{\tau_n'}$ at level $1-\alpha$ as

$$\widetilde{I}^{(2)}_{\tau_n'}(\alpha) = \widetilde{I}^{(2)}_{\tau_n'}(\alpha;J) = \left[\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'} \times \exp\left(\pm\sqrt{\frac{\widetilde{\sigma}_n^2(J)}{v_n(1-k_n/n,\tau_n')}}\, z_{1-\alpha/2}\right)\right]$$

where $\widetilde{\sigma}_n^2(J)$ is obtained from $\sigma_n^2(J)$ by replacing γ by $\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n}$, $q_{1-k_n/n}$ by $\widehat{q}_{1-k_n/n}$ and $\mathbb{E}(Y)$ by $\overline{Y}_n$ in the quantities $V_{11}(J) \equiv \gamma^2$, $V_{12}(J)$, $V_{22}(J)$, $\partial_1 h_n(\gamma,1)$ and $\partial_2 h_n(\gamma,1)$, thus producing their respective estimators $\widehat{V}_{11}(J) \equiv (\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n})^2$, $\widehat{V}_{12}(J)$, $\widehat{V}_{22}(J)$, $\partial_1\widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)$ and $\partial_2\widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)$.

Next, we examine the asymptotic properties of this confidence interval that can be calculated by the R function `CIextExpect` (with `method="indirect"`) available as part of our package `Expectrem`.

Theorem 2. *Assume that $\mathbb{E}|\min(Y,0)| < \infty$ and condition $\mathcal{C}_2(\gamma,\rho,A)$ holds with $0<\gamma<1$, $\rho<0$ and $A(t) = b\gamma t^\rho$. Let k_n, τ_n' be two sequences such that $k_n\to\infty$, $k_n/n\to 0$, $n(1-\tau_n')/k_n \to 0$ and $\log(k_n/(n(1-\tau_n')))/\sqrt{k_n} \to 0$ as $n\to\infty$. Assume further that $\sqrt{k_n}A(n/k_n)\to 0$ and $\sqrt{k_n}/q_{1-k_n/n}\to\lambda\in\mathbb{R}$, and $\overline{\rho}$ and $\overline{b}$ are consistent estimators of ρ and b such that $(\overline{\rho}-\rho)\log(n) = \mathrm{o}_{\mathbb{P}}(1)$. Then:*

(i) For any $J\geq 1$, $\widetilde{\sigma}_n^2(J)\to\gamma^2$ in probability, and

$$\forall\alpha\in(0,1),\ \lim_{n\to\infty}\mathbb{P}\left(\xi_{\tau_n'}\in\widetilde{I}^{(2)}_{\tau_n'}(\alpha;J)\right) = 1-\alpha.$$

(ii) One has

$$\frac{\xi_{\tau_n'}}{\sqrt{k_n}}(\mathrm{length}(\widetilde{I}^{(2)}_{\tau_n'}(\alpha;J)) - \mathrm{length}(\widetilde{I}^{(1)}_{\tau_n'}(\alpha))) \to 2\gamma m(\gamma) z_{1-\alpha/2}$$

in probability, where $m(\gamma) = (1-\gamma)^{-1} - \log(\gamma^{-1}-1)$.

The corrected interval $\widetilde{I}^{(2)}_{\tau_n'}(\alpha;J)$ is not always longer than the naive Gaussian interval $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, since $m(\gamma)>0$ if and only if $\gamma>\gamma_0\approx 0.218$. Interestingly, even when $\widetilde{I}^{(2)}_{\tau_n'}(\alpha;J)$ is shorter than $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, the former tends to have better coverage than the latter in finite samples; this is due to the fact that coverage of $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$ appears to be higher than the nominal level when γ is small. We note moreover that, since $\gamma^3(1-\gamma)/(1-2\gamma)>\gamma^2$ if and only if $\gamma\in((3-\sqrt{5})/2,1/2)\approx(0.382,0.5)$, the corrected interval $\widehat{I}^{(2)}_{\tau_n'}(\alpha;J)$ will tend to be shorter than $\widetilde{I}^{(2)}_{\tau_n'}(\alpha;J)$ when the underlying distribution has a finite third moment. Of course, this rule-of-thumb ignores the variability of the LAWS extreme expectile estimator, which may be considerable in finite samples. We shall check these insights on simulated data in Section 5.

5 Finite-sample study

5.1 The `Expectrem` package

The R package `Expectrem`, freely available at https://github.com/AntoineUC/Expectrem, has been updated to include the methodology developed in this article and fixes have been made to ensure correctness of the implementation. Aside from containing the functions `tindexp` (to compute either $\widehat{\gamma}^E_{\tau_n}$ if argument `br=FALSE` or $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$ if `br=TRUE`) and `extExpect` to compute LAWS (if argument `method="direct"`) or quantile-based (if `method="indirect"`) extreme expectile estimators $\widehat{\xi}^{\star}_{\tau_n'}$ and $\widetilde{\xi}^{\star}_{\tau_n'}$ (the bias-reduced versions $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ and $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ are obtained with `br=TRUE`), we added two functions in order to make the simulation study and applications introduced in this paper easily reproducible:

- `CIextExpect`: an improved version of `extExpect`, which returns the extreme expectile estimate, and its associated bias-reduced and variance-corrected confidence interval $\widehat{I}^{(2)}_{\tau'_n}(\alpha; J)$ (for LAWS estimates) and $\widetilde{I}^{(2)}_{\tau'_n}(\alpha; J)$ (for quantile-based estimates). The required inputs are the dataset `X`, the value $k_n = n(1-\tau_n)$ as `k` and the target expectile level τ'_n as `tau`. The LAWS (`method="direct"`, default) and quantile-based (`method="indirect"`) Weissman estimators can be computed, and the nominal level (`ci.level`, default `0.95`) may be chosen. For comparison purposes, one may also compute the naive, uncorrected confidence intervals $\widehat{I}^{(1)}_{\tau'_n}(\alpha)$ and $\widetilde{I}^{(1)}_{\tau'_n}(\alpha)$, available respectively as `method="direct_naive"` and `method="indirect_naive"`, and univariate versions of the corrected confidence regions of Padoan and Stupfler (2022), available as `method="direct_PS"` and `method="indirect_PS"` for LAWS and quantile-based estimators, respectively.

- `logspacqqplot`: for a dataset `X` given in input, this function returns an exponential quantile-quantile plot of the k_n (as `k`) top log-spacings. If argument `weighted=TRUE` (default), the weighted log-spacings $i \log(X_{n-i+1,n}/X_{n-i,n})$ are computed ($1 \le i \le k$). If `weighted=FALSE`, they are replaced by $\log(X_{n-i+1,n}/X_{n-k_n,n})$. The straight line with slope $\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n}$ may be added if `add.line=TRUE`. This is useful for the purpose of diagnosing tail heaviness *a posteriori* once a choice of k_n has been made.

Note that the expression of $\widetilde{\varphi}^{(2)}_n(\widehat{\xi}_{\tau_n})$, introduced in Section 3 in the calculation of $\widehat{I}^{(2)}_{\tau'_n}(\alpha; J)$, can be somewhat unstable in small samples due to the estimation of γ and ρ. The Cauchy-Schwarz inequality yields $(\varphi^{(1)}(\xi_{\tau_n}))^2 \le \varphi^{(2)}(\xi_{\tau_n}) \le \sqrt{\varphi^{(4)}(\xi_{\tau_n})}$ and motivates the constrained estimator

$$\breve{\varphi}^{(2)}_n(\widehat{\xi}_{\tau_n}) = \min\left\{\max\left\{\widetilde{\varphi}^{(2)}_n(\widehat{\xi}_{\tau_n}), (\widehat{\varphi}^{(1)}_n(\widehat{\xi}_{\tau_n}))^2\right\}, \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{\xi}_{\tau_n})^4 \mathbb{1}\{Y_i > \widehat{\xi}_{\tau_n}\}}\right\}$$

which we found to work reasonably well in practice in a wide array of situations. This constrained estimator is used in the `CIextExpect` function (if `method="direct"`) in place of $\widetilde{\varphi}^{(2)}_n(\widehat{\xi}_{\tau_n})$.

5.2 Simulation study

In order to get a comprehensive overview of the practical performance of the suggested confidence intervals, we consider the following test cases for the distribution of Y (see Table C2 for more details):

- A Fréchet distribution with tail index $\gamma > 0$, *i.e.* $\overline{F}(y) = 1 - \exp(-y^{-1/\gamma})$ for $y > 0$. Here $A(t) = \gamma/(2t)$, and in particular $\rho = -1$. The Fréchet distribution is ubiquitous in extreme value analysis, especially in multivariate problems where it is the standard marginal distribution when estimating multivariate extreme dependence structures, see for instance Chapter 8 of Beirlant et al. (2004).
- A Burr distribution with tail index $\gamma > 0$ and second-order parameter $\rho < 0$, *i.e.* $\overline{F}(y) = (1 + y^{-\rho/\gamma})^{1/\rho}$ for $y > 0$. Here $A(t) = \gamma t^{\rho}$, so we can make ρ vary in order to generate scenarios with various degrees of difficulty. We consider here $\rho = -5, -1, -0.5$, representing easy, standard and hard scenarios, respectively.

For each of these families of distributions, we consider $\gamma = 0.1, 0.2, 0.3$ and 0.4 for the LAWS estimator, and $\gamma = 0.1, 0.3, 0.5$ and 0.7 for the quantile-based estimator: recall that the latter is valid on the larger range of values $\gamma \in (0,1)$, compared to the former which is restricted to $\gamma \in (0, 1/2)$. In an attempt to evaluate further the influence of finite-sample bias upon the performance of each inference technique, we also consider the following two models:

- A Student distribution with number of degrees of freedom $\nu = 10/3$, *i.e.* $\gamma = 1/\nu = 0.3$. Here
$$A(t) = \frac{\nu+1}{\nu+2}\left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)\nu^{(\nu-1)/2}}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\right)^{-2/\nu} t^{-2/\nu} \approx 0.635\, t^{-0.6},$$
and in particular $\rho = -2/\nu = -0.6$.
- A Student distribution with number of degrees of freedom $\nu = 10/3$ and location parameter 1, *i.e.* obtained by adding 1 to a random draw from the previous model. Here again $\gamma = 1/\nu = 0.3$, but by Lemma 1 in

Appendix C.1,

$$A(t) = -\frac{1}{\nu}\left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)\nu^{(\nu-1)/2}}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\right)t^{-1/\nu} \approx -0.265\,t^{-0.3},$$

and therefore $\rho = -1/\nu = -0.3$.
The rationale for considering these last two models is that while half of the Student realizations will be positive on average, we observed that 80% of the observations generated by the shifted Student model were positive, so that theoretically the techniques considered in this paper will be applicable on a larger range of values of k in the last model compared to the vanilla Student model; this should, however, be at the expense of a (theoretically) larger bias, since the decay of the function A to 0 is slower in the shifted Student model.

In each of the considered models, we simulate $N = 5{,}000$ datasets $\{Y_1, \ldots, Y_n\}$ of $n \in \{200, 1{,}000, 5{,}000\}$ independent realizations of Y, with distribution function $F = 1 - \overline{F}$. We set $\tau_n' = 1 - 1/n$ and compare the finite-sample performance of the following inference methods for $\xi_{\tau_n'}$:

- The standard confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$. In order to show the impact of bias reduction, the versions of these intervals not featuring bias correction at all, *i.e.* where $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$, $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$, $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$ and $\widehat{\gamma}^{\mathrm{H,BR}}_{\tau_n}$ are replaced by $\widehat{\xi}^{\star}_{\tau_n'}$, $\widetilde{\xi}^{\star}_{\tau_n'}$, $\widehat{\gamma}^{\mathrm{E}}_{\tau_n}$ and $\widehat{\gamma}^{\mathrm{H}}_{\tau_n}$, respectively, are also considered, namely

$$\widehat{I}^{(3)}_{\tau_n'}(\alpha) = \left[\widehat{\xi}^{\star}_{\tau_n'}\exp\left(\pm\sqrt{\frac{\widehat{s}^2_n}{v_n(\tau_n,\tau_n')}}z_{1-\alpha/2}\right)\right]$$

$$\text{with } \widehat{s}^2_n = \frac{(\widehat{\gamma}^{\mathrm{E}}_{\tau_n})^3(1-\widehat{\gamma}^{\mathrm{E}}_{\tau_n})}{1-2\widehat{\gamma}^{\mathrm{E}}_{\tau_n}}, \tag{5.1}$$

and

$$\widetilde{I}^{(3)}_{\tau_n'}(\alpha) = \left[\widetilde{\xi}^{\star}_{\tau_n'}\exp\left(\pm\sqrt{\frac{\widetilde{\sigma}^2_n}{v_n(\tau_n,\tau_n')}}z_{1-\alpha/2}\right)\right]$$

$$\text{with } \widetilde{\sigma}^2_n = (\widehat{\gamma}^{\mathrm{H}}_{\tau_n})^2. \tag{5.2}$$

- Our proposed confidence intervals $\widehat{I}^{(2)}_{\tau_n'}(\alpha; J)$ and $\widetilde{I}^{(2)}_{\tau_n'}(\alpha; J)$, with $J = 1$ and $J = 2$, respectively.

- The corrected confidence intervals introduced in Padoan and Stupfler (2022), *i.e.* univariate versions of their confidence regions $\widetilde{\mathcal{E}}^{\star}_{\tau_n',\alpha}$ and $\widehat{\mathcal{E}}^{\star}_{\tau_n',\alpha}$ for LAWS and quantile-based estimators, respectively. These take the form

$$\widehat{I}^{(4)}_{\tau_n'}(\alpha) = \left[\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}\exp\left(\pm\sqrt{\frac{\overline{s}^2_n}{v_n(\tau_n,\tau_n')}}z_{1-\alpha/2}\right)\right] \tag{5.3}$$

and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha) =$

$$\left[\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}\exp\left(\pm\sqrt{\frac{\overline{\sigma}^2_n}{v_n(\tau_n,\tau_n')}}\times z_{1-\alpha/2}\right)\right] \tag{5.4}$$

where $\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (resp. $\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$) and $\overline{s}^2_n$ (resp. $\overline{\sigma}^2_n$) are suitably chosen bias-corrected LAWS (resp. quantile-based) estimates of $\xi_{\tau_n'}$ and asymptotic variance estimates for these quantities. We refer to Padoan and Stupfler (2022) for full details.

These are all calculated using the `R` function `CIextExpect` from the `Expectrem` package, described in Section 5.1, at the default $1-\alpha = 0.95$ confidence level.

To assess the finite-sample performance of these inference methods, we provide three types of graphs:

- A graph comparing the empirical variances (across the $N = 5{,}000$ replicated samples and on the log-scale) of $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ (resp. $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$) with the average of the estimated variances $\widehat{s}^{2,\mathrm{BR}}_n$, $\widehat{s}^2_n(J)$, $\widehat{s}^2_n$ and $\overline{s}^2_n$ (resp. $\widetilde{\sigma}^{2,\mathrm{BR}}_n$, $\widetilde{\sigma}^2_n(2)$, $\widetilde{\sigma}^2_n$ and $\overline{\sigma}^2_n$) giving rise to the considered confidence intervals,
- A graph comparing the empirical coverage probability (across the $N = 5{,}000$ replicated samples) of each of the confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha; 1)$, $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widehat{I}^{(4)}_{\tau_n'}(\alpha)$ (resp. $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(2)}_{\tau_n'}(\alpha; 2)$, $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha)$) to the nominal level $1 - \alpha = 0.95$,
- A graph comparing the median length (across the $N = 5{,}000$ replicated samples) of each of these confidence intervals.

This results in 96 series of 3 graphs, for the Fréchet and Burr models, and 12 series of 3 graphs, for the (location) Student model. These are deferred to Appendix D.1, as Figures D1–D28, for the sake of brevity. We give representative examples of our results in Figures 3 and 4.

As the left panels make clear, the confidence intervals $\widehat{I}_{\tau_n'}^{(2)}(\alpha;1)$ and $\widetilde{I}_{\tau_n'}^{(2)}(\alpha;2)$ are constructed based on variance estimates whose median tends to be much closer to the empirical variance of $\widehat{\xi}_{\tau_n'}^{\star,\mathrm{BR}}$ and $\widetilde{\xi}_{\tau_n'}^{\star,\mathrm{BR}}$ than those on which the examined competitors are built. That the average value of our variance estimates may be poor should not be surprising since we are working with heavy-tailed data, whose high values are very disruptive to an accurate estimation of the variance. The middle panels indicate that overall there is substantial gain in preferring $\widehat{I}_{\tau_n'}^{(2)}(\alpha;1)$ (resp. $\widetilde{I}_{\tau_n'}^{(2)}(\alpha;2)$) to $\widehat{I}_{\tau_n'}^{(1)}(\alpha)$ and $\widehat{I}_{\tau_n'}^{(3)}(\alpha)$ (resp. $\widetilde{I}_{\tau_n'}^{(1)}(\alpha)$ and $\widetilde{I}_{\tau_n'}^{(3)}(\alpha)$). Empirical coverage probabilities of $\widehat{I}_{\tau_n'}^{(2)}(\alpha;1)$ and $\widetilde{I}_{\tau_n'}^{(2)}(\alpha;2)$ are also typically more stable than those of $\widehat{I}_{\tau_n'}^{(4)}(\alpha)$ and $\widetilde{I}_{\tau_n'}^{(4)}(\alpha)$ as a function of k, while generally being as close or even closer to the nominal level (see *e.g.* the example of the Burr distribution with $\gamma = 0.3$ and $\rho = -1$, with the LAWS approach, or the difficult case of the Burr distribution with $\gamma = 0.1$ and $\rho = -0.5$, with the quantile-based method). The right panels confirm the observations made following Theorems 1 and 2: the LAWS (resp. quantile-based) interval $\widehat{I}_{\tau_n'}^{(2)}(\alpha;1)$ (resp. $\widetilde{I}_{\tau_n'}^{(2)}(\alpha;2)$) is longer than its counterpart $\widehat{I}_{\tau_n'}^{(1)}(\alpha)$ (resp. longer than $\widetilde{I}_{\tau_n'}^{(1)}(\alpha)$ for values of $\gamma \geq 0.3$), while being kept to a reasonable length. There does not seem to be a rule for comparing these lengths with those of $\widehat{I}_{\tau_n'}^{(4)}(\alpha)$ and $\widetilde{I}_{\tau_n'}^{(4)}(\alpha)$. Finally, in the Student case, which should be considered difficult because the second-order parameter ρ is then close to 0, it is interesting to note that performance is not worse on the shifted distribution compared to the vanilla Student setting. In fact, at least for small k, our variance estimation approach seems to be excellent, while accurately estimating the variance of the (bias-reduced) LAWS estimator for a standard Student sample appears to be very difficult. This is quite surprising, since the second-order auxiliary function $A : t \mapsto -0.265\, t^{-0.3}$ of the shifted Student distribution decays more slowly to 0 than the corresponding function $A : t \mapsto 0.635\, t^{-0.6}$ for the standard Student distribution. We believe that this is likely due to the constant 0.635, in front of the leading power, being greater in absolute value than -0.265, meaning that while the bias should be theoretically expected to be lower for the Student distribution in very large samples, this is not necessarily the case in smaller samples. This last set of results should be viewed as an illustration of the fact that our bias and variance corrections, and indeed our estimation methods generally, are not location-invariant.

5.3 Real data analyses

We showcase the proposed inference methods on three real data examples: the first two revisit the applications of Daouia et al. (2018) and Girard et al. (2022) to medical insurance data and stock price data, respectively, and the third one addresses cyber insurance of data breaches through the assessment of, and inference about, their associated extreme risk.

5.3.1 Example 1: Society of Actuaries medical insurance claims data

The Society of Actuaries (SOA Group) Medical Insurance Large Claims Database records all the claim amounts exceeding 25,000 USD over the timeframe 1991-92. We only deal here with the $n = 75{,}789$ claims for 1991. This dataset (available in the R package `Expectrem` as the dataset `soa`) has been analyzed by a number of authors, including Beirlant et al. (2004) who find evidence of a heavy right tail (see also the histogram and rug plot in the top left panel of Figure 5) with tail index $\gamma \approx 1/3$, and Daouia et al. (2018), who estimate extreme quantiles at level $\alpha_n = 1 - 1/100{,}000$ by making use of extrapolated extreme expectile estimates.

Our objective here is to estimate and infer extreme expectiles themselves in this dataset using our bias-reduced and variance-corrected techniques. The top right panel of Figure 5 compares the point estimates $\widehat{\xi}_{\tau_n'}^{\star}$ in (2.1) and $\widetilde{\xi}_{\tau_n'}^{\star}$ in (2.2) with their bias-reduced counterparts in (2.5) and (2.6), as functions of the intermediate anchor level $k = k_n = n(1-\tau_n)$ and at level $\tau_n' = 1 - 1/100{,}000$. It is readily seen that the bias reduction scheme is highly effective, which here stems from the fact that the estimated value of ρ hovers around $\overline{\rho} = -0.2$ and thus the bias component is expected to be sizeable. We then

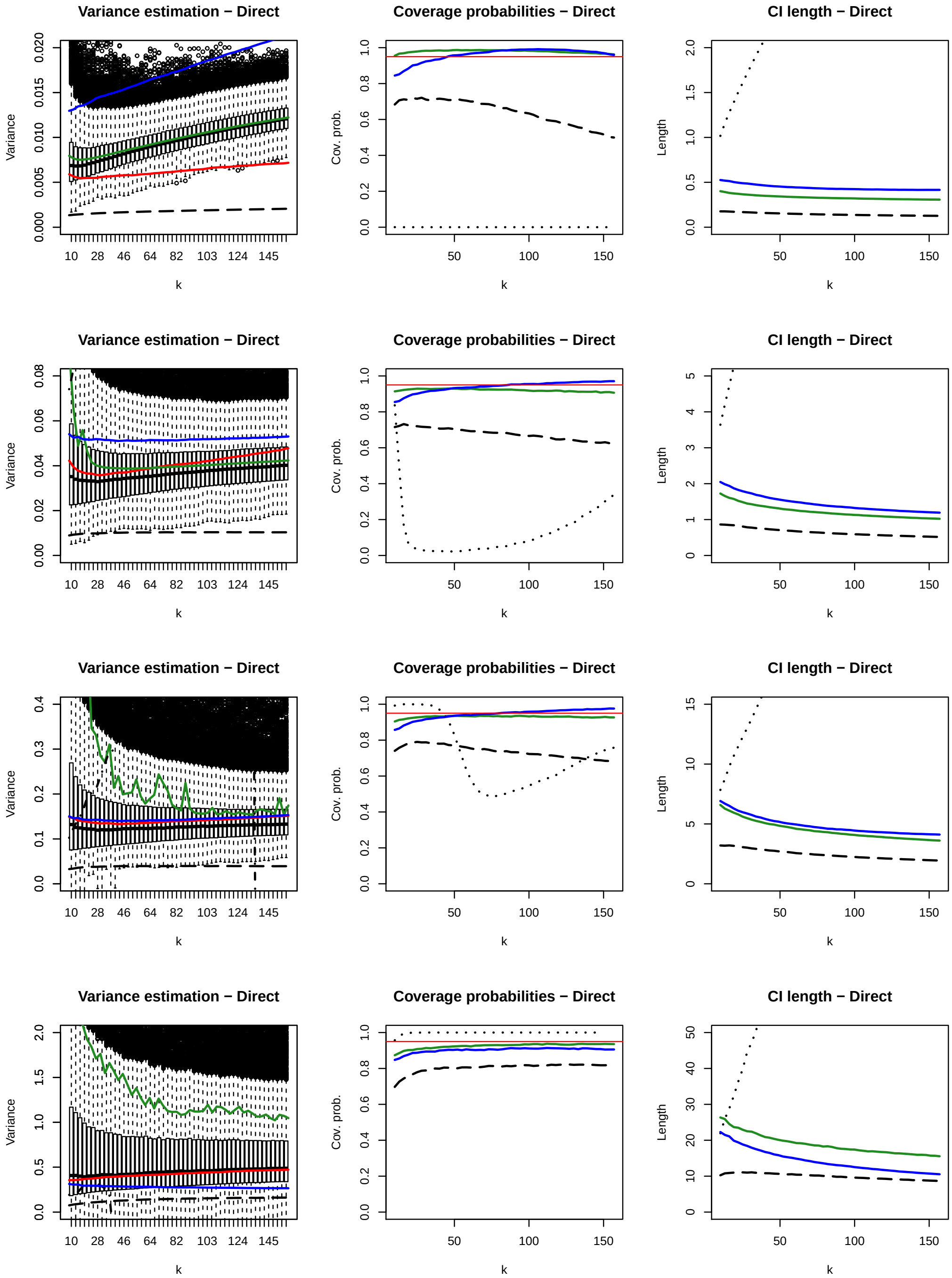

Fig. 3: Left panels: Comparison of the empirical variance of $\sqrt{v_n(\tau_n,\tau_n')}\log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widehat{s}^2_n(1)$ and the average values of $\widehat{s}^{2,\mathrm{BR}}_n$ (dashed curve), $\widehat{s}^2_n(1)$ (green curve), $\widetilde{s}^2_n$ (dashed-dotted curve) and $\overline{s}^2_n$ (blue curve) as functions of $k=k_n=n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha;1)$, $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widehat{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). First row: Burr distribution with $\gamma=0.1$ and $\rho=-0.5$, second row: Fréchet distribution with $\gamma=0.2$, third row: Burr distribution with $\gamma=0.3$ and $\rho=-1$, fourth row: Burr distribution with $\gamma=0.4$ and $\rho=-5$. The sample size is $n=1{,}000$ and the target expectile level is $\tau_n'=1-1/n=0.999$.

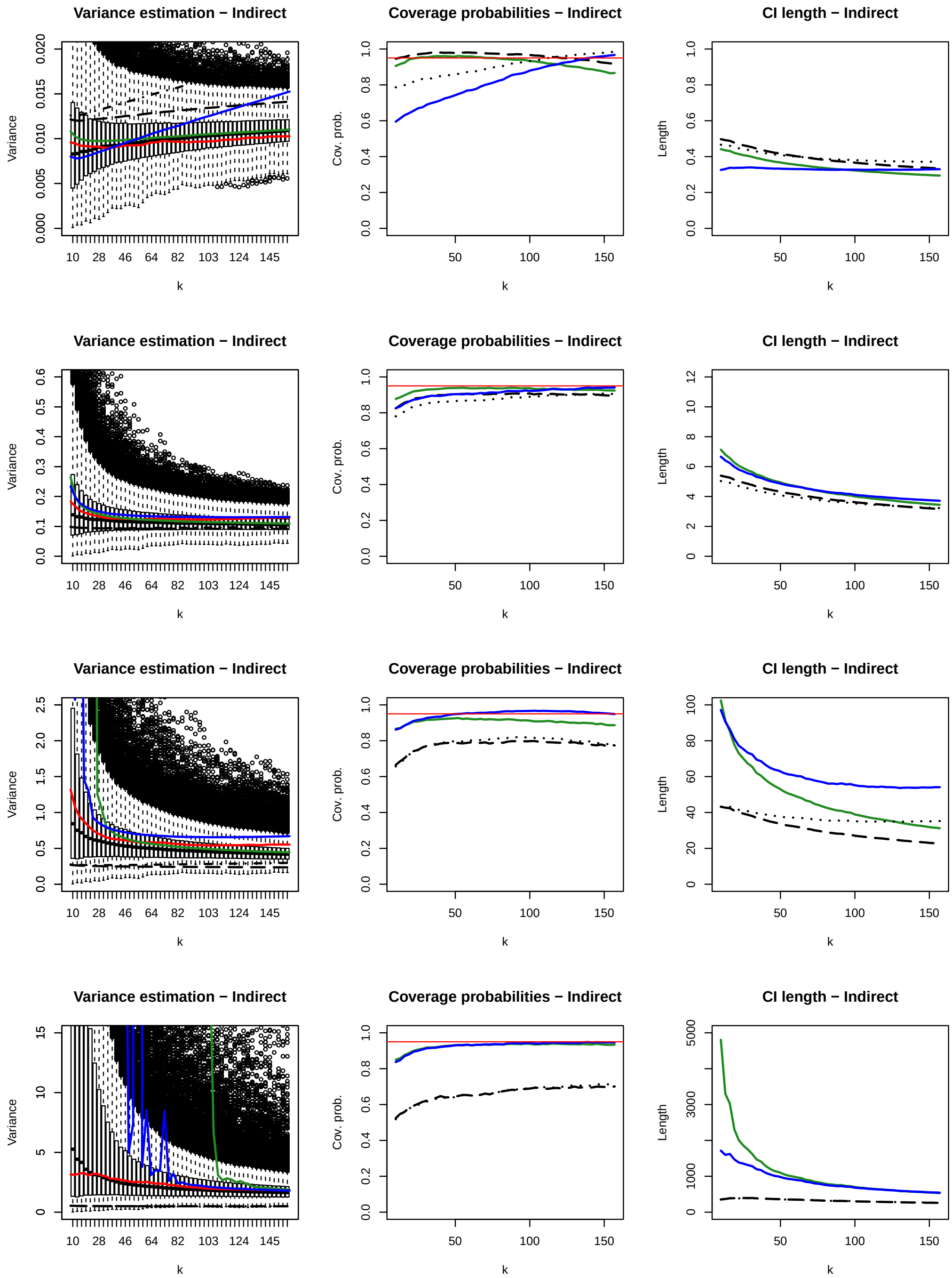

Fig. 4: Left panels: Comparison of the empirical variance of $\sqrt{v_n(\tau_n,\tau_n')}\log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widetilde{\sigma}_n^2(2)$ and the average values of $\widetilde{\sigma}_n^{2,\mathrm{BR}}$ (dashed curve), $\widetilde{\sigma}_n^2(2)$ (green curve), $\widetilde{\sigma}_n^2$ (dashed-dotted curve) and $\overline{\sigma}_n^2$ (blue curve) as functions of $k=k_n=n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(2)}_{\tau_n'}(\alpha;2)$, $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). First row: Burr distribution with $\gamma=0.1$ and $\rho=-0.5$, second row: Fréchet distribution with $\gamma=0.3$, third row: Burr distribution with $\gamma=0.5$ and $\rho=-1$, fourth row: Burr distribution with $\gamma=0.7$ and $\rho=-5$. The sample size is $n=1{,}000$ and the target expectile level is $\tau_n'=1-1/n=0.999$.

carry out inference on the basis of the bias-reduced estimates, see the bottom panels of Figure 5, where we compare the straightforward intervals $\widehat{I}^{(1)}_{\tau'_n}(\alpha)$ and $\widetilde{I}^{(1)}_{\tau'_n}(\alpha)$ to the variance-corrected versions $\widehat{I}^{(2)}_{\tau'_n}(\alpha;1)$ and $\widetilde{I}^{(2)}_{\tau'_n}(\alpha;2)$, respectively. We first observe that the variance-corrected interval is very similar to its naive counterpart when the quantile-based method is used. This is sensible in view of the results of our simulation study, where one sees that the naive and corrected quantile-based extreme expectile confidence intervals are often close when $\gamma \approx 0.3$, which is precisely the case here. When using the LAWS-based confidence intervals, however, the bottom left panel of Figure 5 suggests that the finite-sample variability of the LAWS extreme expectile estimate is more important than what is suggested by the straightforward interval $\widehat{I}^{(1)}_{\tau'_n}(\alpha)$. Using the latter would lead to a substantial underestimation of the uncertainty on the target extreme expectile and thus to an incorrect assessment of tail risk carried out from large medical claims, with potentially detrimental consequences on the insurance companies involved.

5.3.2 Example 2: Commerzbank stock price data

We consider the daily negative log-returns of the Commerzbank stock price on the DAX30 stock exchange between March 6, 2012 and July 28, 2016, resulting in a sample $Y_1, \dots, Y_n$ of size $n =$ 1,048 (available in the R package `Expectrem` as the dataset `commerzbank`). This dataset is naturally a time series with substantial serial dependence, so that we cannot apply the bias and variance corrections we develop to the raw data; nevertheless, and as already pointed out by Girard et al. (2022) for this dataset, the temporal dependence can be handled in a dynamic extreme value estimation setup by, first, filtering the time series with an ARMA(1,1)-GARCH(1,1) model:

$$Y_t = \mu + \phi Y_{t-1} + u_t + \theta u_{t-1},$$

where $u_t = \sigma_t \varepsilon_t$ is such that $\sigma_t^2 = \mathfrak{c} + \mathfrak{a} u_{t-1}^2 + \mathfrak{b}\sigma_{t-1}^2$, with (ε_t) being an unobserved independent nondegenerate white noise sequence, *i.e.* copies of a random variable ε such that $\mathbb{E}(\varepsilon) = 0$, $\mathbb{E}(\varepsilon^2) = 1$ and $\mathbb{P}(\varepsilon^2 = 1) < 1$, and the constants μ, ϕ, θ, $\mathfrak{a}$, $\mathfrak{b}$ and $\mathfrak{c}$ being the model parameters. Denote by $\mathcal{F}_n$ the algebra generated by the ARMA-GARCH process up to time n, and recall that the expectile is positive homogeneous and location equivariant, so that the conditional expectile for the next day given the data up to time n can be modeled as

$$\xi_\tau(Y_{n+1}|\mathcal{F}_n) = \mu + \phi Y_n + \sigma_{n+1}\, \xi_\tau(\varepsilon) + \theta u_n.$$

The estimation of expectiles in practice requires an informed choice of the target level. One possible strategy, studied among others by Bellini and Di Bernardino (2017), is to assume that the underlying distribution function Φ of the innovations ε_t is standard Gaussian. For example, in this setting, an expectile of level $\tau'_n = 0.995$ of the ε_t coincides with their quantile at level $\beta_n = \Phi(\xi_{\tau'_n}(\varepsilon)) \approx 0.974$. As such, if the predicted expectiles and quantiles at the levels $\tau'_n = 0.995$ and 0.974, respectively, are significantly different, then there is evidence that the distribution of the errors is not Gaussian.

We first estimate the ARMA-GARCH model parameters and predict the residuals $\widehat{\varepsilon}_i$ and $\widehat{u}_i$ using the methodology introduced in Girard et al. (2022), *i.e.* by, first, using the function `garchFit` in the `R` package `fGarch` and the option `@residuals` in order to retrieve the $\widehat{u}_i$, and then, by obtaining the predictions $\widehat{\varepsilon}_i$ of the innovations by fitting a pure GARCH model directly to the $\widehat{u}_i$ and applying `garchFit(...)@residuals`. Following the theory developed in Girard et al. (2021), we treat the residuals $\widehat{\varepsilon}_i$ from the model as independent and identically distributed copies of ε for the estimation of the tail index γ of ε and its extreme expectile $\xi_{\tau'_n}(\varepsilon)$. The independence assumption was checked using a series of Ljung-Box independence tests on residuals and their squares, with the lowest $p-$value being 0.39. Evidence that ε is indeed heavy-tailed is gathered in the top-right panel of Figure 6 using exponential QQ-plots of the log-spacings. The extreme ($\tau'_n =$ 0.995) expectile estimates with their associated 90% and 95% confidence intervals (as functions of $k = k_n = n(1-\tau_n)$) are reported in the middle panels of Figure 6. We eyeball the threshold k_n for stability of the estimates and take in the sequel $k_n = 32$, leading to the tail index estimates $\widehat{\gamma}^{\mathrm{H,BR}}_{k_n}(\varepsilon) = 0.329$ and $\widehat{\gamma}^{\mathrm{E,BR}}_{k_n}(\varepsilon) = 0.315$. The extreme quantile (of level β_n, extrapolated and non-extrapolated) estimators are also reported in

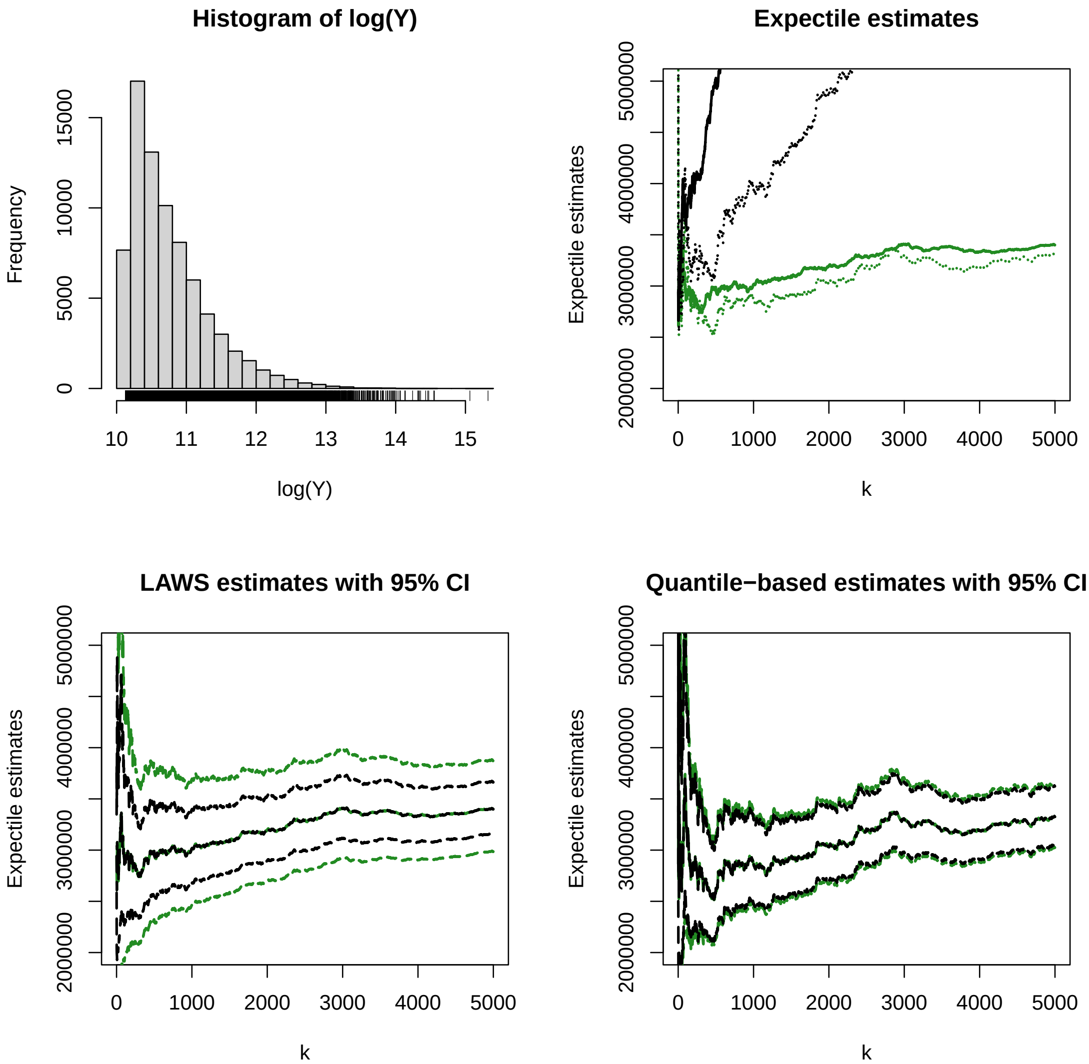

Fig. 5: Society of Actuaries medical insurance claims data (sample size $n = 75{,}789$). Top left: Histogram and rug plot of the log-claims $\log Y_i$, $1 \le i \le n$. Top right: Expectile point estimates $\widehat{\xi}^{\star}_{\tau'_n}$ (solid black), $\widetilde{\xi}^{\star}_{\tau'_n}$ (dotted black), $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}$ (solid green) and $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}$ (dotted green), at level $\tau'_n = 1 - 1/100{,}000$, as a function of $k = k_n = n(1-\tau_n)$. Bottom left: Bias-reduced LAWS expectile point estimate $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}$ (solid black) with associated confidence intervals $\widehat{I}^{(1)}_{\tau'_n}(\alpha)$ (dashed black) and $\widehat{I}^{(2)}_{\tau'_n}(\alpha;1)$ (dashed green) at the 95% confidence level. Bottom right: Bias-reduced quantile-based expectile point estimate $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}$ (solid black) with associated confidence intervals $\widetilde{I}^{(1)}_{\tau'_n}(\alpha)$ (dashed black) and $\widetilde{I}^{(2)}_{\tau'_n}(\alpha;2)$ (dashed green) at the 95% confidence level.

the middle panels of Figure 6, where

$$\widehat{q}_{\beta_n}(\varepsilon) = \widehat{\varepsilon}_{n-\lfloor n(1-\beta_n)\rfloor,n} \text{ and}$$
$$\widehat{q}^{\star,\mathrm{BR}}_{\beta_n}(\varepsilon) = \left(\frac{k_n}{n(1-\beta_n)}\right)^{\widehat{\gamma}^{\mathrm{H,BR}}_{k_n}(\varepsilon)} \widehat{\varepsilon}_{n-k_n,n}$$
$$\times \left(1+\frac{1}{\overline{\rho}}\left\{\left(\frac{k_n}{n(1-\beta_n)}\right)^{\overline{\rho}}-1\right\}\overline{b}\widehat{\gamma}^{\mathrm{H,BR}}_{k_n}(\varepsilon)\left(\frac{n}{k_n}\right)^{\overline{\rho}}\right)$$

are the empirical and bias-reduced extrapolated Weissman quantile estimates of $q_{\beta_n}(\varepsilon)$. The resulting quantile and expectile estimates are reported, with confidence intervals, in Table 1 for $k_n = 32$.

Dynamic predictions of extreme expectiles of Y_{n+1} given Y_n are then obtained as

$$\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(Y_{n+1}|\mathcal{F}_n) = \widehat{\mu} + \widehat{\phi}Y_n + \widehat{\sigma}_{n+1}\,\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\varepsilon) + \widehat{\theta}\widehat{u}_n,$$
$$\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(Y_{n+1}|\mathcal{F}_n) = \widehat{\mu} + \widehat{\phi}Y_n + \widehat{\sigma}_{n+1}\,\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\varepsilon) + \widehat{\theta}\widehat{u}_n.$$

In order to visualize these predictions, the key point is that, if $\widehat{\mu}$, $\widehat{\phi}$, $\widehat{\theta}$ and the past $(Y_t)_{t\le n-1}$ of the process are considered as fixed, then $\widehat{u}_n = Y_n - \widehat{\phi}Y_{n-1} - \widehat{\theta}\widehat{u}_{n-1} - \widehat{\mu}$ (resp. $\widehat{\sigma}^2_{n+1}$) is an affine function of Y_n (resp. is quadratic in $\widehat{u}_n$ and hence in Y_n). The dynamic extreme expectile estimates can thus be seen as functions of the observation $Y_n = y_n$. These functions are represented in the bottom panels of Figure 6, together with confidence intervals deduced from those of $\xi_{\tau_n'}(\varepsilon)$ by ignoring the uncertainty in $\widehat{\mu}$, $\widehat{\phi}$, $\widehat{\theta}$, $\widehat{\mathfrak{a}}$, $\widehat{\mathfrak{b}}$ and $\widehat{\mathfrak{c}}$: this is justified theoretically by the fact that these estimators converge at the rate $1/\sqrt{n}$ and thus much faster than those obtained via our extreme value procedures. They are compared with dynamic bias-reduced Weissman quantile estimates of level β_n, calculated as

$$\widehat{q}^{\star,\mathrm{BR}}_{\beta_n}(Y_{n+1}|\mathcal{F}_n) = \widehat{\mu} + \widehat{\phi}Y_n + \widehat{\sigma}_{n+1}\,\widehat{q}^{\star,\mathrm{BR}}_{\beta_n}(\varepsilon) + \widehat{\theta}\widehat{u}_n.$$

It is readily seen that the quantile estimates do not belong to the 90% confidence intervals of the expectile estimates, and belong to the LAWS-based 95%−confidence interval but not to the quantile-based 95%−confidence interval. This constitutes further evidence that the residuals of the ARMA-GARCH model are not Gaussian and therefore that the standard selection rules of expectile levels on light-tailed data do not apply to stock price data in general. A financial consequence of this analysis is that a risk assessment based on expectiles using the guidelines provided by Bellini and Di Bernardino (2017) in the Gaussian case would be more conservative than if it were based on quantiles.

5.3.3 Example 3: Privacy Rights Clearinghouse cyber risk data

Cyber risks, especially the subclass of personal data breaches from firms/organisations, defined by Article 4(12) of the European Union General Data Protection Regulation as breaches of security leading to mass identity fraud, are a rapidly increasing threat to individuals, companies, public services, and governments. Insurance against data breaches constitutes the bulk of coverage of losses related to cyber events. The financial loss due to a single breached piece of private data is estimated to be 213 USD on average, which translates into a global cost estimated at hundreds of billions of USD per year, see Wheatley et al. (2016). Early work on data breach risks from an extreme value perspective includes quantifying the heavy-tailed nature of breach sizes in Maillart and Sornette (2010) and Wheatley et al. (2016). Very recently, Farkas et al. (2021) have computed loss quantiles for a cyber portfolio by combining Generalized Pareto modeling and regression trees. We further investigate such questions here using expectiles. As described below in our exploratory analysis, tail index estimates are found to be mostly larger than 0.5 in the different studied classes of breached data, thus leading to more conservative (*i.e.* prudent) risk measurements against catastrophic breach events than if quantiles were used, according to Bellini et al. (2014).

We use the PRC (Privacy Rights Clearinghouse) database[3] which is the most comprehensive open scientific dataset for breaches occurring in businesses, educational institutions, government and military, healthcare institutions, and other media and nonprofit organisations. The PRC database gathers cyber events from different sources, which introduces heterogeneity among the reported events. In order to reduce the heterogeneity effect when inferring the tail risk, we follow the setup of Farkas et al. (2021) and let the sources of information be grouped into four types:

[3]See https://privacyrights.org/data-breaches. Data on file with the authors.

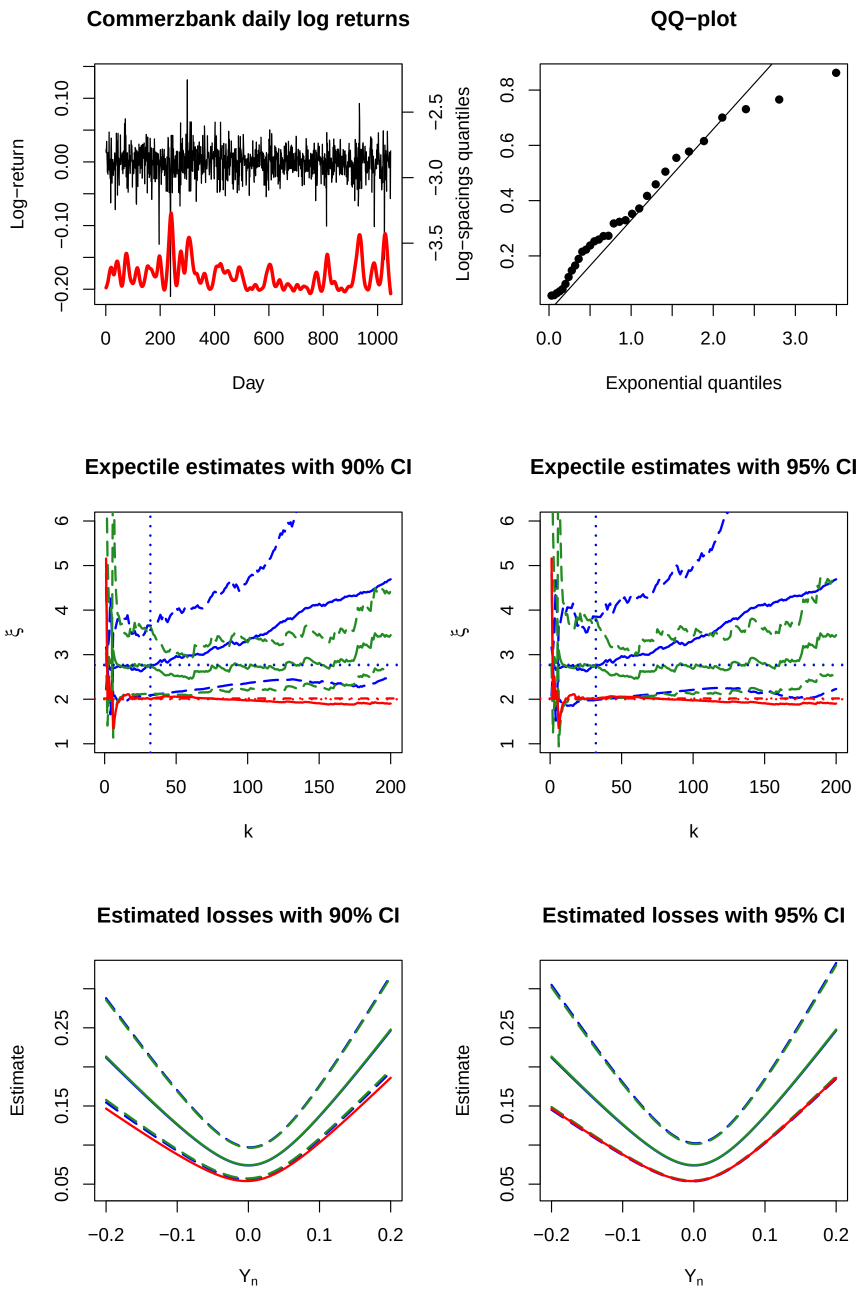

Fig. 6: Commerzbank stock price data (sample size $n = 1{,}048$). Top left: Daily log-returns (black curve) and ARMA-GARCH log-volatility estimates (red bold curve, smoothed using the R function `smooth.spline` with smoothing parameter $\lambda = 5 \times 10^{-7}$). Top right: Exponential QQ-plot of the log-spacings $\log(\widehat{\varepsilon}_{n-i+1,n}/\widehat{\varepsilon}_{n-k_n,n})$, $1 \le i \le k_n = 32$. The straight line has slope $\widehat{\gamma}^{\mathrm{H,BR}}_{k_n}(\varepsilon) = 0.329$. Middle panels: Expectile estimates $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\varepsilon)$ and $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\varepsilon)$ at $\tau'_n = 0.995$ (solid blue and green curves) with their associated confidence intervals $\widehat{I}^{(2)}_{\tau'_n}(\alpha; 1)$ and $\widetilde{I}^{(2)}_{\tau'_n}(\alpha; 2)$ (dashed blue and green curves) at the 90% (left) and 95% (right) confidence levels, along with quantile estimates $\widehat{q}_{\beta_n}(\varepsilon)$ and $\widehat{q}^{\star,\mathrm{BR}}_{\beta_n}(\varepsilon)$ at level $\beta_n \approx 0.974$ (dashed-dotted and solid red curves). Bottom panels: Dynamic estimates and their associated 90% (left) and 95% (right) confidence intervals, with the same color code.

Interval	90% CI	95% CI	$\xi_{\tau_n'}(\varepsilon)$ estimate	$\widehat{q}_{\beta_n}(\varepsilon)$	$\widehat{q}^{\star}_{\beta_n}(\varepsilon)$
$\widehat{I}^{(1)}_{\tau_n'}(\alpha)$	[2.435, 3.136]	[2.377, 3.213]	$\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'} \approx 2.763$	2.017	1.999
$\widehat{I}^{(2)}_{\tau_n'}(\alpha)$	[2.091, 3.652]	[1.982, 3.852]			
$\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$	[2.337, 3.305]	[2.261, 3.416]	$\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'} \approx 2.779$		
$\widetilde{I}^{(2)}_{\tau_n'}(\alpha)$	[2.129, 3.627]	[2.023, 3.817]			

Table 1: 90% (first column) and 95% (second column) confidence intervals for the extreme expectile $\xi_{\tau_n'}$ of the innovations ε_t at $\tau_n' = 0.995$, along with their associated LAWS and quantile-based point estimates (third column), and empirical and Weissman quantile estimates (fourth and fifth columns) calculated at the level $\beta_n \approx 0.974$.

(1) US federal government agencies; (2) US state government agencies; (3) Nonprofit organizations; and (4) Media. To strive towards checking the fundamental assumption of identically distributed data, we stratify the breaches occurring in healthcare, medical providers and medical insurance services (represented by the MED acronym) and educational institutions (represented by the EDU acronym) along the source of information and breach event type. This differentiation of the cyber events (after eliminating duplicates and NA values), occurring between January 1st, 2005 and October 25th, 2019, results in 6 MED (respectively, 3 EDU) classes of sizes $n > 100$; we exclude small clusters of less than 100 observations. As summarized in Table 2, the classes we obtain are determined by the following five types of breaches: HACK (hacking by an outside party or infection by malware), INSD (similar to HACK but performed by an insider), PHYS (physical breach, including paper documents that are lost, discarded or stolen), PORT (portable device lost, discarded or stolen) and DISC (unintended disclosure not involving hacking, intentional breach or physical loss), and by Federal, State and Nonprofit sources of information (the Media source appears only in very small clusters of size $n < 50$).

A more global and automated clustering is adopted by Farkas et al. (2021) based on their GPD regression tree analysis of the severity distribution, but nothing guarantees that our model assumptions (*e.g.* stationarity) are satisfied when using the leaves of their trees. It should also be noted that the PRC database only reports breach sizes, that is, the number of records affected by each breach event. The related financial loss is not provided directly by the database, but can be quantified approximately from a claims-driven loss formula, calibrated by Farkas et al. (2021) to the specific PRC database. We refer to Section 5.1 in Farkas et al. (2021) for a detailed discussion and justification of the rationale behind their calibration. The resulting losses provide a rough approximation of the real claim data. Lack of this data still constitutes a bottleneck to precise risk assessment and insurance pricing, but the analysis presented here might already help "risk thinkers" to better grasp these new cyber risks from the extreme value perspective.

Numerical results obtained for the nine studied classes are displayed in Table 2. The tuning parameter $k = n(1-\tau_n)$ appearing in both $\widehat{\gamma}^{\mathrm{H,BR}}_{\tau_n}$ and $\widetilde{\xi}^{\star,\mathrm{BR}}_{1-1/n}$ is chosen following the path stability algorithm developed by El Methni and Stupfler (2017) which consists in computing the standard deviations of the estimators over a "moving window" of successive values of k, and then by selecting the first value of k where the standard deviation is minimal and sufficiently low for each estimator. Graphs of $\widehat{\gamma}^{\mathrm{H,BR}}_{1-k/n}$ versus k, along with asymptotic Gaussian 95% confidence intervals and the value of k selected by the path stability procedure as well as the final pointwise estimate, and their analogs for $\widetilde{\xi}^{\star,\mathrm{BR}}_{1-1/n}$ can be found in Appendix D.2.

The results gathered in Table 2 show a substantial difference in the bias-corrected Hill estimates $\widehat{\gamma}^{\mathrm{H,BR}}_{\tau_n} \in [0.40, 0.78]$ of the tail index across all clusters. We also observe 7 extremely heavy-tailed cyber clusters with tail index estimates exceeding 0.5 (as indicated in green), in which the estimated expectile risk measure at level $\tau_n' =$

$1-1/n$ exceeds the maximum historical loss (as indicated in purple). Our asymptotic corrected Gaussian confidence intervals point towards upper risk bounds whose value may even exceed three times the maximum loss. An accurate quantification of this uncertainty is of course crucial, especially in this challenging context with small datasets. We provide a further point of comparison by estimating a quantile-based Value-at-Risk using the bias-corrected extrapolated Weissman estimator $\widehat{q}^{\star,\mathrm{BR}}_{1-1/n}$. It is precisely on the 7 extremely heavy-tailed cyber risk classes where the difference between extreme expectiles and quantiles is strongest. In particular, the estimated expectile loss measures can be as much as three times higher than their quantile counterparts, and the quantile estimates do not even exceed the maximum historical loss for five of these classes (as indicated in orange). It is here more prudent to measure tail risk based on expectiles rather than on quantiles, which solely depend on the frequency of tail losses and not on their severity.

Acknowledgments

The authors acknowledge an anonymous reviewer for helpful comments that led to an improved version of this article. This research was supported by the French National Research Agency under the grants ANR-19-CE40-0013 (ExtremReg project), ANR-17-EURE-0010 (EUR CHESS) and ANR-11-LABX-0020-01 (Centre Henri Lebesgue). A. Daouia and G. Stupfler acknowledge financial support from the TSE-HEC ACPR Chair and from an AXA Research Fund Award on "Mitigating risk in the wake of the COVID-19 pandemic".

Statements and Declarations

The authors declare no competing interests. All authors contributed equally to the work, and read and approved the final manuscript.

References

Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Mathematical Finance **9**(3), 203–228 (1999)

Gneiting, T.: Making and evaluating point forecasts. Journal of the American Statistical Association **106**(494), 746–762 (2011)

Bellini, F., Klar, B., Müller, A., Gianin, E.R.: Generalized quantiles as risk measures. Insurance: Mathematics and Economics **54**, 41–48 (2014)

Ziegel, J.F.: Coherence and elicitability. Mathematical Finance **26**(4), 901–918 (2016)

Newey, W.K., Powell, J.L.: Asymmetric least squares estimation and testing. Econometrica **55**(4), 819–847 (1987)

Holzmann, H., Klar, B.: Expectile asymptotics. Electronic Journal of Statistics **10**(2), 2355–2371 (2016)

Krätschmer, V., Zähle, H.: Statistical inference for expectile-based risk measures. Scandinavian Journal of Statistics **44**(2), 425–454 (2017)

Daouia, A., Girard, S., Stupfler, G.: Estimation of tail risk based on extreme expectiles. Journal of the Royal Statistical Society: Series B **80**(2), 263–292 (2018)

Daouia, A., Girard, S., Stupfler, G.: Extreme M-quantiles as risk measures: from L^1 to L^p optimization. Bernoulli **25**(1), 264–309 (2019)

Daouia, A., Girard, S., Stupfler, G.: Tail expectile process and risk assessment. Bernoulli **26**(1), 531–556 (2020)

Padoan, S.A., Stupfler, G.: Joint inference on extreme expectiles for multivariate heavy-tailed distributions. Bernoulli **28**(2), 1021–1048 (2022)

Girard, S., Stupfler, G., Usseglio-Carleve, A.: On automatic bias reduction for extreme expectile estimation. Statistics & Computing **32**(4), 64 (2022)

Athreya, K.B.: Bootstrap of the mean in the infinite variance case. Annals of Statistics **15**(2), 724–731 (1987)

Knight, K.: On the bootstrap of the sample mean in the infinite variance case. Annals of Statistics **17**(3), 1168–1175 (1989)

Sector, level, type of breach	n	$\widehat{\gamma}_{\tau_n}^{\mathrm{H,BR}}$ $(\widehat{k}_{\mathrm{opt}})$ [95% CI]	$\widetilde{\xi}_{1-1/n}^{\star,\mathrm{BR}}$ $(\widetilde{k}_{\mathrm{opt}})$ (million USD) [95% CI]	$\max_{1\leq i\leq n} Y_i$ (million USD)	$\widehat{q}_{1-1/n}^{\star,\mathrm{BR}}$ $(\check{k}_{\mathrm{opt}})$ (million USD) [95% CI]
MED, federal, PHYS	1140	0.654 (320) [0.582, 0.726]	156.04 (322) [79.91, 304.70]	95.12	84.17 (316) [55.49, 127.68]
MED, federal, DISC	766	0.582 (299) [0.516, 0.648]	61.54 (323) [36.26, 104.44]	57.07	33.51 (314) [23.11, 48.58]
MED, federal, HACK	545	0.782 (263) [0.688, 0.877]	545.01 (265) [181.26, 1638.73]	463.34	157.36 (266) [92.96, 266.36]
MED, state, HACK	178	0.497 (67) [0.378, 0.616]	17.49 (69) [9.21, 33.22]	60.26	13.07 (42) [7.21, 23.69]
MED, nonprofit, PORT	223	0.714 (54) [0.523, 0.904]	169.99 (42) [37.90, 762.35]	60.26	99.86 (56) [47.12, 211.64]
MED, nonprofit, INSD	106	0.752 (42) [0.525, 0.980]	38.03 (43) [4.40, 328.57]	16.57	12.18 (40) [5.13, 28.92]
EDU, nonprofit, HACK	178	0.402 (36) [0.270, 0.533]	29.17 (43) [16.57, 51.34]	32.63	28.54 (68) [17.49, 46.56]
EDU, nonprofit, DISC	174	0.529 (27) [0.329, 0.728]	21.48 (27) [7.10, 64.94]	21.13	15.98 (28) [8.53, 29.91]
EDU, nonprofit, PORT	103	0.578 (46) [0.411, 0.745]	25.03 (46) [8.80, 71.17]	10.52	11.78 (58) [6.36, 2181]

Table 2: Privacy Rights Clearinghouse cyber risk data. First column: Sector, level and type of breach. Second column: Sample size n of each class. Third column: Bias-reduced Hill estimate of γ along with the selected k and asymptotic Gaussian 95% confidence interval. Fourth column: Quantile-based bias-reduced extreme expectile estimator $\widetilde{\xi}_{1-1/n}^{\star,\mathrm{BR}}$ along with the selected k and bias-reduced and variance-corrected asymptotic Gaussian 95% confidence interval $\widetilde{I}_{1-1/n}^{(2)}(\alpha;2)$. Fifth column: Maximum observation. Sixth column: Weissman extreme quantile estimator $\widehat{q}_{1-1/n}^{\star,\mathrm{BR}}$ obtained by making use of the bias-reduced Hill estimator, along with the selected k and asymptotic Gaussian 95% confidence interval.

Hall, P.: Asymptotic properties of the bootstrap for heavy-tailed distributions. Annals of Probability **18**(3), 1342–1360 (1990)

Angus, J.: Asymptotic theory for bootstrapping the extremes. Communications in Statistics - Theory and Methods **22**(1), 15–30 (1993)

Hall, P., Jing, B.-Y.: Comparison of bootstrap and asymptotic approximations to the distribution of a heavy-tailed mean. Statistica Sinica **8**(3), 887–906 (1998)

Guillou, A.: Bootstrap confidence intervals for the Pareto index. Communications in Statistics - Theory and Methods **29**(1), 211–226 (2000)

Cornea-Madeira, A., Davidson, R.: A parametric bootstrap for heavy-tailed distributions. Econometric Theory **31**(3), 449–470 (2015)

Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling Extremal Events for Insurance and Finance. Springer, Berlin-Heidelberg (1997)

Bellini, F., Di Bernardino, E.: Risk management with expectiles. The European Journal of Finance **23**(6), 487–506 (2017)

Hill, B.M.: A simple general approach to inference about the tail of a distribution. The Annals of Statistics **3**(5), 1163–1174 (1975)

Girard, S., Stupfler, G., Usseglio-Carleve, A.: Nonparametric extreme conditional expectile estimation. Scandinavian Journal of Statistics **49**(1), 78–115 (2022)

Caeiro, F., Gomes, M.I., Pestana, D.: Direct reduction of bias of the classical Hill estimator. Revstat **3**(2), 113–136 (2005)

Jones, M.C.: Expectiles and M-quantiles are quantiles. Statistics & Probability Letters **20**, 149–153 (1994)

Stupfler, G., Usseglio-Carleve, A.: Composite bias-reduced L^p-quantile-based estimators of extreme quantiles and expectiles. Canadian Journal of Statistics **51**(2), 704–742 (2023)

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

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.: Statistics of Extremes: Theory and Applications. Wiley, Chichester, United Kingdom (2004)

Girard, S., Stupfler, G., Usseglio-Carleve, A.: Extreme conditional expectile estimation in heavy-tailed heteroscedastic regression models. Annals of Statistics **49**(6), 3358–3382 (2021)

Wheatley, S., Maillart, T., Sornette, D.: The extreme risk of personal data breaches and the erosion of privacy. European Physical Journal B **89**(7), 1–12 (2016)

Maillart, T., Sornette, D.: Heavy-tailed distribution of cyber-risks. European Physical Journal B **75**(3), 357–364 (2010)

Farkas, S., Lopez, O., Thomas, M.: Cyber claim analysis using Generalized Pareto regression trees with applications to insurance. Insurance: Mathematics and Economics **98**, 92–105 (2021)

El Methni, J., Stupfler, G.: Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions. Statistica Sinica **27**(2), 907–930 (2017)

Appendix A Estimation of the second-order parameters b and ρ

Under the working assumption that $A(t) = b\gamma t^\rho$, the estimators $\overline{b} = \overline{b}_n$ and $\overline{\rho} = \overline{\rho}_n$ of the b and ρ parameters are calculated as follows: for a given $n-$sample $(Y_1, \ldots, Y_n)$, let $Y_{1,n} \le Y_{2,n} \le \cdots \le Y_{n,n}$ be the associated order statistics and set

$$M_{\kappa_n}^{(j)} = \frac{1}{\kappa_n} \sum_{i=1}^{\kappa_n} \left(\log Y_{n-i+1,n} - \log Y_{n-\kappa_n,n}\right)^j, \text{ for } j = 1, 2, 3.$$

For $j = 1$, this is just the Hill estimator. These quantities are the basic building blocks for the quantity $T_{\kappa_n}^{(\tau)}$ defined as

$$T_{\kappa_n}^{(\tau)} = \begin{cases} \dfrac{\left(M_{\kappa_n}^{(1)}\right)^\tau - \left(M_{\kappa_n}^{(2)}/2\right)^{\tau/2}}{\left(M_{\kappa_n}^{(2)}/2\right)^{\tau/2} - \left(M_{\kappa_n}^{(3)}/6\right)^{\tau/3}} & \text{if } \tau > 0, \\[2ex] \dfrac{\log\left(M_{\kappa_n}^{(1)}\right) - \frac{1}{2}\log\left(M_{\kappa_n}^{(2)}/2\right)}{\frac{1}{2}\log\left(M_{\kappa_n}^{(2)}/2\right) - \frac{1}{3}\log\left(M_{\kappa_n}^{(3)}/6\right)} & \text{if } \tau = 0. \end{cases}$$

The estimator of ρ that we consider is a simple function of $T_{\kappa_n}^{(\tau)}$:

$$\widehat{\rho}_{\kappa_n}^{(\tau)} = -\left|\frac{3(T_{\kappa_n}^{(\tau)} - 1)}{T_{\kappa_n}^{(\tau)} - 3}\right|. \tag{A1}$$

This estimator is implemented in the R function `mop`, available in the `Expectrem` package[4]. In this package, with $\kappa_n = \lfloor n^{0.999} \rfloor$, and a choice of $\tau \in \{0, 1\}$ is made based on a stability criterion for $\kappa \mapsto \widehat{\rho}_\kappa^{(\tau)}$ for large κ. According to Proposition 2.1 in Caeiro et al. (2005), these choices ensure, if ρ is large enough (a calculation analogous to that in Remark 2.2 in Caeiro et al. (2005) provides roughly $\rho > -249.75$, which will cover all practical applications), that $(\widehat{\rho}_{\kappa_n}^{(\tau)} - \rho)\log(n) = \mathrm{o}_{\mathbb{P}}(1)$ as required in our asymptotic results. An estimator of b is then

$$\widehat{b}_{\kappa_n} = \left(\frac{\kappa_n}{n}\right)^{\overline{\rho}} \frac{\left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-\overline{\rho}}\right)\left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n} U_i\right) - \left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-\overline{\rho}} U_i\right)}{\left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-\overline{\rho}}\right)\left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-\overline{\rho}} U_i\right) - \left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-2\overline{\rho}} U_i\right)}, \tag{A2}$$

where $\overline{\rho} = \widehat{\rho}_{\kappa_n}^{(\tau)}$ and the $U_i = i\log(Y_{n-i+1,n}/Y_{n-i,n})$ are the weighted log-spacings. This estimator is also available from the R function `mop`. The aforementioned choice of κ_n ensures that $\overline{b} = \widehat{b}_{\kappa_n}$ is consistent, see Proposition 2.2 in Caeiro et al. (2005). Our results will therefore always feature the conditions that $(\overline{\rho} - \rho)\log(n) = \mathrm{o}_{\mathbb{P}}(1)$ and $\overline{b} - b = \mathrm{o}_{\mathbb{P}}(1)$.

[4]Taken from the package `evt0`, unfortunately not available anymore from CRAN at the time this paper was written.

Appendix B Detailed calculations for the construction of LAWS and quantile-based extreme expectile confidence intervals

B.1 LAWS confidence interval construction

Our argument is motivated by an inspection of the existing proof of the joint asymptotic normality of an intermediate empirical expectile and the corresponding expectile-based tail index estimator $\widehat{\gamma}^{\mathrm{E}}_{\tau_n}$ (see Theorem 2 in Stupfler and Usseglio-Carleve (2023) and Theorem 2 in Girard et al. (2022)). The idea of the proof is not new; however, carefully identifying the crucial steps of this proof and the key approximations that are made will be instrumental in our subsequent construction of refined confidence intervals in which the novelty of the present work resides. We assume throughout this section that $\mathbb{E}|\min(Y,0)|^2 < \infty$ and $\gamma < 1/2$.

B.1.1 Preliminary steps

The basic idea of the proof behind Theorem 2 in Stupfler and Usseglio-Carleve (2023) is that for any $z_1, z_2 \in \mathbb{R}$,

$$\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} - (\gamma^{-1}-1)\right) \le z_1,\ \sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \le z_2\right\}$$
$$= \left\{\widehat{\xi}_{\tau_n} \ge \widehat{q}_{1-(1-\tau_n)(\gamma^{-1}-1+z_1/\sqrt{n(1-\tau_n)})},\ \sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \le z_2\right\}. \tag{B1}$$

Evaluating the joint uncertainty in $(\widehat{\gamma}^{\mathrm{E}}_{\tau_n}, \widehat{\xi}_{\tau_n})$ is therefore equivalent to evaluating the joint uncertainty in $(\widehat{\xi}_{\tau_n}, \widehat{q}_{\alpha_n})$, where α_n is a sequence satisfying $1-\alpha_n = (\gamma^{-1}-1)(1-\tau_n)(1+\mathrm{o}(1))$. Pick then $u, v \in \mathbb{R}$ and set

$$A_n(u,v) = \left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \le u\right\} \cap \left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}} - 1\right) \le v\right\}.$$

Setting $y_n = y_n(u) = \xi_{\tau_n}(1 + u/\sqrt{n(1-\tau_n)})$ and $z_n = z_n(v) = q_{\alpha_n}(1 + v/\sqrt{n(1-\tau_n)})$, it follows that

$$A_n(u,v) = \left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{E}}_n(y_n)}{\overline{E}(y_n)} - 1\right) \le \sqrt{n(1-\tau_n)}\left(\frac{\overline{E}(\xi_{\tau_n})}{\overline{E}(y_n)} - 1\right)\right\}$$
$$\cap \left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(z_n)}{\overline{F}(z_n)} - 1\right) \le \sqrt{n(1-\tau_n)}\left(\frac{\overline{F}(q_{\alpha_n})}{\overline{F}(z_n)} - 1\right)\right\}. \tag{B2}$$

There are two crucial points of note here. On the one hand, it is straightforward to obtain, via a Lyapunov central limit theorem, that

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{E}}_n(y_n)}{\overline{E}(y_n)} - 1, \frac{\widehat{\overline{F}}_n(z_n)}{\overline{F}(z_n)} - 1\right) \xrightarrow{d} \mathcal{N}\left((0,0), \frac{\gamma}{1-\gamma}\boldsymbol{\Sigma}\right), \tag{B3}$$

where the 2×2 symmetric matrix $\boldsymbol{\Sigma}$ has entries $\Sigma_{11} = 2(1-\gamma)/(1-2\gamma)$, $\Sigma_{12} = \Sigma_{22} = 1$. On the other hand,

$$\sqrt{n(1-\tau_n)} \left(\frac{\overline{E}(\xi_{\tau_n})}{\overline{E}(y_n)} - 1 \right) = \frac{u}{\gamma}(1 + \mathrm{o}(1)) \tag{B4}$$
$$\text{and } \sqrt{n(1-\tau_n)} \left(\frac{\overline{F}(q_{\alpha_n})}{\overline{F}(z_n)} - 1 \right) = \frac{v}{\gamma}(1 + \mathrm{o}(1)).$$

This is shown using the second-order regular variation property and Lemma A.3(iv) in Stupfler and Usseglio-Carleve (2023). One now readily concludes that

$$\sqrt{n(1-\tau_n)} \left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1, \frac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}} - 1 \right) \xrightarrow{d} \mathcal{N}\left((0,0), \frac{\gamma^3}{1-\gamma} \boldsymbol{\Sigma} \right). \tag{B5}$$

The final step is to recall Equation (B1) and to use the fact that

$$\begin{aligned} &\{\widehat{\xi}_{\tau_n} \geq \widehat{q}_{\alpha_n}\} \\ &= \left\{ \sqrt{n(1-\tau_n)} \left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) \geq \sqrt{n(1-\tau_n)} \left(\frac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}} - 1 \right) \frac{q_{\alpha_n}}{\xi_{\tau_n}} + \sqrt{n(1-\tau_n)} \left(\frac{q_{\alpha_n}}{\xi_{\tau_n}} - 1 \right) \right\}. \end{aligned}$$

Now recall that, from Proposition 1 in Daouia et al. (2020) and the second-order regular variation assumption,

$$\begin{aligned} \frac{q_{\alpha_n}}{\xi_{\tau_n}} = \frac{q_{\alpha_n}}{q_{\tau_n}} \times \frac{q_{\tau_n}}{\xi_{\tau_n}} &= \left(1 - \frac{\gamma^2}{1-\gamma} \times \frac{z_1}{\sqrt{n(1-\tau_n)}} (1 + \mathrm{o}(1)) \right) \\ &\times \left(1 - \frac{\gamma(\gamma^{-1}-1)^{\gamma}}{q_{\tau_n}} (\mathbb{E}(Y) + \mathrm{o}(1)) - \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} + \mathrm{o}(1) \right) A((1-\tau_n)^{-1}) \right) \end{aligned}$$

when $\alpha_n = \alpha_n(z_1) = 1 - (1-\tau_n)(\gamma^{-1} - 1 + z_1/\sqrt{n(1-\tau_n)})$. Neglecting bias terms and isolating z_1, one now readily finds, *via* (B1), that

$$\begin{aligned} &\left\{ \sqrt{n(1-\tau_n)} \left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} - (\gamma^{-1} - 1) \right) \leq z_1, \ \sqrt{n(1-\tau_n)} \left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) \leq z_2 \right\} \\ &= \left\{ -\frac{1-\gamma}{\gamma^2} \times \sqrt{n(1-\tau_n)} \left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) + \frac{1-\gamma}{\gamma^2} \times \sqrt{n(1-\tau_n)} \left(\frac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}} - 1 \right) + \mathrm{o}_{\mathbb{P}}(1) \leq z_1, \right. \\ &\qquad \left. \sqrt{n(1-\tau_n)} \left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) \leq z_2 \right\}. \end{aligned}$$

This means that the asymptotic covariance matrix of

$$\sqrt{n(1-\tau_n)} \left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} - (\gamma^{-1} - 1), \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right)$$

is essentially that of

$$\begin{pmatrix} -\dfrac{1-\gamma}{\gamma^2} & \dfrac{1-\gamma}{\gamma^2} \\ 1 & 0 \end{pmatrix} \times \sqrt{n(1-\tau_n)} \begin{pmatrix} \dfrac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \\ \dfrac{\widehat{q}_{1-(1-\tau_n)(\gamma^{-1}-1)}}{q_{1-(1-\tau_n)(\gamma^{-1}-1)}} - 1 \end{pmatrix}. \tag{B6}$$

The asymptotic covariance structure of $\sqrt{n(1-\tau_n)}(\widehat{\gamma}^{\mathrm{E}}_{\tau_n} - \gamma, \widehat{\xi}_{\tau_n}/\xi_{\tau_n} - 1)$ is then obtained by applying the delta-method with the function $x \mapsto (h(x), y) = (1/(1+x), y)$: one finds

$$\sqrt{n(1-\tau_n)} \left(\widehat{\gamma}^{\mathrm{E}}_{\tau_n} - \gamma, \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right)$$
$$\xrightarrow{d} \mathcal{N}\left(\left(\frac{\gamma(\gamma^{-1}-1)^{1-\rho}}{1-\gamma-\rho}\lambda_1 + \gamma^2(\gamma^{-1}-1)^{\gamma+1}\mathbb{E}(Y)\lambda_2, \, 0 \right), \frac{\gamma^3}{1-2\gamma} \begin{pmatrix} 1-\gamma & 1 \\ 1 & 2 \end{pmatrix} \right) \tag{B7}$$

where $\lambda_1 = \lim_{n\to\infty} \sqrt{n(1-\tau_n)} A((1-\tau_n)^{-1})$ and $\lambda_2 = \lim_{n\to\infty} \sqrt{n(1-\tau_n)}/q_{\tau_n}$, assumed henceforth to be finite. Recall finally that, if $(1-\tau_n')/(1-\tau_n) \to 0$,

$$\left(\frac{1-\tau_n'}{1-\tau_n} \right)^{\gamma} \frac{\xi_{\tau_n'}}{\xi_{\tau_n}} = 1 + \mathrm{O}(1/\sqrt{n(1-\tau_n)}),$$

see Proposition 1 in Daouia et al. (2020) and the proof of Theorem 4.3.8 p.139 in de Haan and Ferreira (2006). Using the identity

$$\log\left(\frac{\widehat{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}} \right) = (\widehat{\gamma}^{\mathrm{E}}_{\tau_n} - \gamma) \log\left(\frac{1-\tau_n}{1-\tau_n'} \right) + \log\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} \right) - \log\left(\left(\frac{1-\tau_n'}{1-\tau_n} \right)^{\gamma} \frac{\xi_{\tau_n'}}{\xi_{\tau_n}} \right)$$

provides (2.3). The fact that $\overline{b} - b = \mathrm{o}_{\mathbb{P}}(1)$ and $(\overline{\rho} - \rho)\log(n) = \mathrm{o}_{\mathbb{P}}(1)$ then yields the same asymptotic normality result for $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$, only with the asymptotic mean being 0.

B.1.2 Calculations

The idea is to construct versions of steps (B3), (B4), (B5), (B6) and (B7) that are more accurate in finite samples. We do so based on the sequence $\alpha_n = 1 - \overline{F}(\xi_{\tau_n})$, which is asymptotically equivalent (but not equal) to the sequence $1 - (1-\tau_n)(\gamma^{-1}-1)$ corresponding to the choice $z_1 = 0$ in (B1). This choice is motivated by the fact that for the construction of asymptotic Gaussian confidence regions, one is typically interested in an accurate calculation of the probability of events such as (B1) for small values of (z_1, z_2), because this is the region where most of the mass of the Gaussian limiting distribution concentrates. Taking $\alpha_n = 1 - \overline{F}(\xi_{\tau_n})$, rather than $\alpha_n = 1 - (1-\tau_n)(\gamma^{-1}-1)$, makes it possible to more accurately quantify the uncertainty in $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$ by comparing it directly to its population counterpart $\overline{F}(\xi_{\tau_n})$.

We first examine what room for improvement there is when using the asymptotic approximation (B3). We do so with y_n replaced by ξ_{τ_n} and z_n replaced by q_{α_n}, *i.e.* we take $u = v = 0$ in $y_n = y_n(u)$ and $z_n = z_n(v)$, which is once again appropriate for our purpose of correctly evaluating the probability of $A_n(u,v)$ for small u and v. It is straightforward to prove that the covariance matrix of

$$\sqrt{n(1-\tau_n)} \left(\frac{\widehat{\varphi}^{(1)}_n(\xi_{\tau_n})}{\varphi^{(1)}(\xi_{\tau_n})} - 1, \frac{\widehat{\overline{F}}_n(q_{\alpha_n})}{\overline{F}(q_{\alpha_n})} - 1 \right)$$

is *exactly* the 2×2 symmetric matrix $\mathbf{M}^\varphi$, where $\mathbf{M}^\varphi = \mathbf{M}_n^\varphi$ has components:

$$M_{n,11}^\varphi = (1-\tau_n)\left(\frac{\varphi^{(2)}(\xi_{\tau_n})}{[\varphi^{(1)}(\xi_{\tau_n})]^2} - 1\right),$$
$$M_{n,12}^\varphi = (1-\tau_n)\left(\frac{\varphi^{(1)}(\xi_{\tau_n} \vee q_{\alpha_n}) + (\xi_{\tau_n} \vee q_{\alpha_n} - \xi_{\tau_n})\overline{F}(\xi_{\tau_n} \vee q_{\alpha_n})}{\varphi^{(1)}(\xi_{\tau_n})(1-\alpha_n)} - 1\right)$$
$$= (1-\tau_n)\frac{\alpha_n}{1-\alpha_n} \text{ and } M_{n,22}^\varphi = (1-\tau_n)\frac{\alpha_n}{1-\alpha_n} = M_{12}^\varphi.$$

In particular

$$\mathbf{M}_n^\varphi \to \begin{pmatrix} \dfrac{2\gamma}{1-2\gamma} & \dfrac{\gamma}{1-\gamma} \\ \dfrac{\gamma}{1-\gamma} & \dfrac{\gamma}{1-\gamma} \end{pmatrix} \text{ as } n \to \infty.$$

Recall that $\widehat{\overline{E}}_n(y) = \widehat{\varphi}_n^{(1)}(y)/(2\widehat{\varphi}_n^{(1)}(y) + y - \overline{Y}_n)$ and therefore

$$\frac{\widehat{\overline{E}}_n(\xi_{\tau_n})}{\overline{E}(\xi_{\tau_n})} - 1$$
$$= \frac{\widehat{\varphi}_n^{(1)}(\xi_{\tau_n})(\xi_{\tau_n} - \mathbb{E}(Y)) - \varphi^{(1)}(\xi_{\tau_n})(\xi_{\tau_n} - \overline{Y}_n)}{\varphi^{(1)}(\xi_{\tau_n})(2\widehat{\varphi}_n^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \overline{Y}_n)}$$
$$= \frac{(\widehat{\varphi}_n^{(1)}(\xi_{\tau_n}) - \varphi^{(1)}(\xi_{\tau_n}))(\xi_{\tau_n} - \mathbb{E}(Y)) + \varphi^{(1)}(\xi_{\tau_n})(\overline{Y}_n - \mathbb{E}(Y))}{\varphi^{(1)}(\xi_{\tau_n})(2\widehat{\varphi}_n^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \overline{Y}_n)}$$
$$\approx \frac{\xi_{\tau_n} - \mathbb{E}(Y)}{2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y)}\left(\frac{\widehat{\varphi}_n^{(1)}(\xi_{\tau_n})}{\varphi^{(1)}(\xi_{\tau_n})} - 1\right) + \frac{1}{2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y)}(\overline{Y}_n - \mathbb{E}(Y)).$$

Since $\overline{Y}_n - \mathbb{E}(Y)$ converges to 0 at the rate $1/\sqrt{n}$ by the central limit theorem, this identity suggests that it is in fact reasonable to write

$$\frac{\widehat{\overline{E}}_n(\xi_{\tau_n})}{\overline{E}(\xi_{\tau_n})} - 1 \approx \frac{\xi_{\tau_n} - \mathbb{E}(Y)}{2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y)}\left(\frac{\widehat{\varphi}_n^{(1)}(\xi_{\tau_n})}{\varphi^{(1)}(\xi_{\tau_n})} - 1\right).$$

As a consequence, the joint distribution of

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{E}}_n(\xi_{\tau_n})}{\overline{E}(\xi_{\tau_n})} - 1, \frac{\widehat{\overline{F}}_n(q_{\alpha_n})}{\overline{F}(q_{\alpha_n})} - 1\right)$$

has a covariance matrix that can be more accurately approximated by $\mathbf{M}^E = \mathbf{M}_n^E$ having components:

$$M_{n,11}^E = \frac{(\xi_{\tau_n} - \mathbb{E}(Y))^2 M_{11}^\varphi}{(2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y))^2},\ M_{n,12}^E = \frac{(\xi_{\tau_n} - \mathbb{E}(Y)) M_{12}^\varphi}{2\varphi^{(1)}(\xi_{\tau_n}) + \xi_{\tau_n} - \mathbb{E}(Y)},\ M_{n,22}^E = M_{22}^\varphi.$$

[Clearly $\mathbf{M}_n^E$ is elementwise asymptotically equivalent to $\mathbf{M}_n^\varphi$.] We now focus on obtaining a more accurate version of (B5). Recall from (B2) that

$$\mathbb{P}(A_n(u,v)) = \mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{E}}_n(y_n)}{\overline{E}(y_n)} - 1\right) \le \sqrt{n(1-\tau_n)}\left(\frac{\overline{E}(\xi_{\tau_n})}{\overline{E}(y_n)} - 1\right)\right\}\right.$$

$$\cap\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(z_n)}{\overline{F}(z_n)}-1\right)\le\sqrt{n(1-\tau_n)}\left(\frac{\overline{F}(q_{\alpha_n})}{\overline{F}(z_n)}-1\right)\right\}\Bigg)$$
$$\approx\mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{E}}_n(\xi_{\tau_n})}{\overline{E}(\xi_{\tau_n})}-1\right)\le\sqrt{n(1-\tau_n)}\left(\frac{\overline{E}(\xi_{\tau_n})}{\overline{E}(y_n)}-1\right)\right\}\right.$$
$$\left.\cap\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(q_{\alpha_n})}{\overline{F}(q_{\alpha_n})}-1\right)\le\sqrt{n(1-\tau_n)}\left(\frac{\overline{F}(q_{\alpha_n})}{\overline{F}(z_n)}-1\right)\right\}\right)$$

where $y_n=y_n(u)=\xi_{\tau_n}(1+u/\sqrt{n(1-\tau_n)})$ and $z_n=z_n(v)=q_{\alpha_n}(1+v/\sqrt{n(1-\tau_n)})$; the approximation is motivated by the fact that $\widehat{\overline{E}}_n(y_n)/\overline{E}(y_n)-1$ and $\widehat{\overline{E}}_n(\xi_{\tau_n})/\overline{E}(\xi_{\tau_n})-1$ have the same asymptotic behavior, and similarly for $\widehat{\overline{F}}_n(z_n)/\overline{F}(z_n)-1$ and $\widehat{\overline{F}}_n(q_{\alpha_n})/\overline{F}(q_{\alpha_n})-1$. Note also that the function $\overline{E}$ is absolutely continuous, because $y\mapsto\varphi^{(1)}(y)=\int_y^\infty\overline{F}(t)\,dt$ is Lipschitz continuous and the denominator $y\mapsto\mathbb{E}(|Y-y|)=2\varphi^{(1)}(y)+y-\mathbb{E}(Y)$ defines a Lipschitz continuous function that is bounded away from zero. The function $\overline{E}$ has Lebesgue derivative

$$\overline{E}'(y)=-\frac{\varphi^{(1)}(y)+\overline{F}(y)(y-\mathbb{E}(Y))}{(2\varphi^{(1)}(y)+y-\mathbb{E}(Y))^2}.$$

This suggests the use of a Taylor expansion to obtain the following more precise version of (B4):

$$\frac{\overline{E}(\xi_{\tau_n})}{\overline{E}(y_n)}-1=u\xi_{\tau_n}\frac{1}{\sqrt{n(1-\tau_n)}}\frac{\varphi^{(1)}(\xi_{\tau_n})+\overline{F}(\xi_{\tau_n})(\xi_{\tau_n}-\mathbb{E}(Y))}{\varphi^{(1)}(\xi_{\tau_n})(2\varphi^{(1)}(\xi_{\tau_n})+\xi_{\tau_n}-\mathbb{E}(Y))}(1+\mathrm{o}(1)).$$

Therefore, using (B2), the distribution of $(\widehat{\xi}_{\tau_n}/\xi_{\tau_n}-1,\widehat{q}_{\alpha_n}/q_{\alpha_n}-1)$ is approximated by that of

$$\left(\frac{1}{\xi_{\tau_n}}\frac{\varphi^{(1)}(\xi_{\tau_n})(2\varphi^{(1)}(\xi_{\tau_n})+\xi_{\tau_n}-\mathbb{E}(Y))}{\varphi^{(1)}(\xi_{\tau_n})+\overline{F}(\xi_{\tau_n})(\xi_{\tau_n}-\mathbb{E}(Y))}\left(\frac{\widehat{\overline{E}}_n(\xi_{\tau_n})}{\overline{E}(\xi_{\tau_n})}-1\right),\gamma\left(\frac{\widehat{\overline{F}}_n(q_{\alpha_n})}{\overline{F}(q_{\alpha_n})}-1\right)\right).$$

An approximation of the covariance matrix of $\sqrt{n(1-\tau_n)}(\widehat{\xi}_{\tau_n}/\xi_{\tau_n}-1,\widehat{q}_{\alpha_n}/q_{\alpha_n}-1)$ is then given by $\mathbf{M}^\xi=\mathbf{M}^\xi_n$, whose elements are

$$M^\xi_{n,11}=\left(\frac{\varphi^{(1)}(\xi_{\tau_n})(2\varphi^{(1)}(\xi_{\tau_n})+\xi_{\tau_n}-\mathbb{E}(Y))}{\xi_{\tau_n}(\varphi^{(1)}(\xi_{\tau_n})+\overline{F}(\xi_{\tau_n})(\xi_{\tau_n}-\mathbb{E}(Y)))}\right)^2M^E_{n,11},$$
$$M^\xi_{n,12}=\gamma\left(\frac{\varphi^{(1)}(\xi_{\tau_n})(2\varphi^{(1)}(\xi_{\tau_n})+\xi_{\tau_n}-\mathbb{E}(Y))}{\xi_{\tau_n}(\varphi^{(1)}(\xi_{\tau_n})+\overline{F}(\xi_{\tau_n})(\xi_{\tau_n}-\mathbb{E}(Y)))}\right)M^E_{n,12},\ \text{and}\ M^\xi_{n,22}=\gamma^2M^E_{n,22}.$$

Note that indeed

$$\mathbf{M}^\xi_n\to\frac{\gamma^3}{1-\gamma}\mathbf{\Sigma}=\begin{pmatrix}\frac{2\gamma^3}{1-2\gamma}&\frac{\gamma^3}{1-\gamma}\\\frac{\gamma^3}{1-\gamma}&\frac{\gamma^3}{1-\gamma}\end{pmatrix}\ \text{as}\ n\to\infty,$$

as a straightforward consequence of Lemma 3 in Stupfler and Usseglio-Carleve (2023) involving burdensome calculations, which use the asymptotic equivalents

$$\varphi^{(1)}(y)=\frac{\gamma}{1-\gamma}y\overline{F}(y)(1+\mathrm{o}(1))\ \text{and}\ \varphi^{(2)}(y)=\frac{2\gamma^2}{(1-\gamma)(1-2\gamma)}y^2\overline{F}(y)(1+\mathrm{o}(1)),$$

combined with the convergence $\overline{F}(\xi_\tau)/(1-\tau) \to \gamma^{-1}-1$ as $\tau \uparrow 1$. This means that $\mathbf{M}^\xi$ provides a corrected version of the asymptotic covariance matrix of $\sqrt{n(1-\tau_n)}(\widehat{\xi}_{\tau_n}/\xi_{\tau_n}-1, \widehat{q}_{\alpha_n}/q_{\alpha_n}-1)$ that is consistent with the asymptotics in (B5).

We turn to finding an improvement of (B6). Note that

$$\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}-\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right)\le z\right\} = \left\{\widehat{\xi}_{\tau_n} \ge \widehat{q}_{1-(1-\tau_n)(\overline{F}(\xi_{\tau_n})/(1-\tau_n)+z/\sqrt{n(1-\tau_n)})}\right\}.$$

The intermediate quantile level $\beta_n = \beta_n(z) = 1-(1-\tau_n)(\overline{F}(\xi_{\tau_n})/(1-\tau_n)+z/\sqrt{n(1-\tau_n)})$ in the right-hand side is such that $1-\beta_n \sim \overline{F}(\xi_{\tau_n}) \sim (\gamma^{-1}-1)(1-\tau_n)$. Then

$$\begin{aligned}
&\mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}-\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right)\le z\right\}\cap\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)\le z'\right\}\right)\\
&=\mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)\ge\sqrt{n(1-\tau_n)}\left(\frac{\widehat{q}_{\beta_n}}{q_{\beta_n}}-1\right)\frac{q_{\beta_n}}{\xi_{\tau_n}}+\sqrt{n(1-\tau_n)}\left(\frac{q_{\beta_n}}{\xi_{\tau_n}}-1\right)\right\}\right.\\
&\qquad\left.\cap\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)\le z'\right\}\right)\\
&\approx\mathbb{P}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}}-1\right)\frac{q_{\alpha_n}}{\xi_{\tau_n}}-\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)\le-\sqrt{n(1-\tau_n)}\left(\frac{q_{\beta_n}}{\xi_{\tau_n}}-1\right)\right\}\right.\\
&\qquad\left.\cap\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)\le z'\right\}\right)
\end{aligned}$$

where again $\alpha_n = \beta_n(0) = F(\xi_{\tau_n})$, because $(\widehat{q}_{\beta_n}/q_{\beta_n}-1)q_{\beta_n}/\xi_{\tau_n}$ and $(\widehat{q}_{\alpha_n}/q_{\alpha_n}-1)q_{\alpha_n}/\xi_{\tau_n}$ have the same asymptotic behavior. From Proposition 1 in Daouia et al. (2020) and the second-order regular variation assumption,

$$\sqrt{n(1-\tau_n)}\left(\frac{q_{\beta_n}}{\xi_{\tau_n}}-1\right)=\sqrt{n(1-\tau_n)}\left(\frac{q_{\beta_n}}{q_{\tau_n}}\times\frac{q_{\tau_n}}{\xi_{\tau_n}}-1\right)\approx-\gamma\frac{1-\tau_n}{\overline{F}(\xi_{\tau_n})}z$$

up to bias terms, which will be taken care of in the bias reduction procedure (and therefore can be neglected at this stage). Conclude that the asymptotic distribution of

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}-\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}, \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)$$

can be more accurately approximated by that of

$$\begin{pmatrix} -\dfrac{1}{\gamma}\dfrac{\overline{F}(\xi_{\tau_n})}{1-\tau_n} & \dfrac{1}{\gamma}\dfrac{\overline{F}(\xi_{\tau_n})}{1-\tau_n} \\ 1 & 0 \end{pmatrix}\sqrt{n(1-\tau_n)}\begin{pmatrix}\dfrac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1 \\ \dfrac{\widehat{q}_{\alpha_n}}{q_{\alpha_n}}-1\end{pmatrix}.$$

As a consequence, this asymptotic distribution can be approximated by a centered Gaussian distribution having covariance matrix $\mathbf{M}=\mathbf{M}_n$ defined by its elements

$$M_{n,11}=\frac{1}{\gamma^2}\left(\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right)^2\left[M^\xi_{n,11}-2M^\xi_{n,12}+M^\xi_{n,22}\right],$$

$$M_{n,12} = \frac{1}{\gamma}\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\left[M^{\xi}_{n,12} - M^{\xi}_{n,11}\right] \text{ and } M_{n,22} = M^{\xi}_{n,11}.$$

This is the required refinement of (B6), satisfying

$$\mathbf{M}_n = \begin{pmatrix} M_{n,11} & M_{n,12} \\ M_{n,12} & M_{n,22} \end{pmatrix} \to \begin{pmatrix} \dfrac{1-\gamma}{\gamma(1-2\gamma)} & -\dfrac{\gamma}{1-2\gamma} \\ -\dfrac{\gamma}{1-2\gamma} & \dfrac{2\gamma^3}{1-2\gamma} \end{pmatrix} \text{ as } n \to \infty.$$

It remains to find a more precise version of (B7) tailored to $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$. This convergence relies on applying the delta-method to $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n)$ with the transformation $\psi : x \mapsto 1/(1+x)$. An even more accurate approximation is possible by, first, noting that $\widehat{\gamma}^{\mathrm{E}}_{\tau_n} = \psi(\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))$ is an (asymptotically) unbiased estimator of $\psi(\overline{F}(\xi_{\tau_n})/(1-\tau_n))$, and then by writing a formal power series expansion of ψ in order to assess the behavior of $\widehat{\gamma}^{\mathrm{E}}_{\tau_n} - \psi(\overline{F}(\xi_{\tau_n})/(1-\tau_n))$. This is done as follows:

$$\begin{aligned}
&\sqrt{n(1-\tau_n)}(\widehat{\gamma}^{\mathrm{E}}_{\tau_n} - \psi(\overline{F}(\xi_{\tau_n})/(1-\tau_n))) \\
&= \sqrt{n(1-\tau_n)}(\psi(\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n)) - \psi(\overline{F}(\xi_{\tau_n})/(1-\tau_n))) \\
&= \sum_{j=1}^{\infty} \frac{\psi^{(j)}(\overline{F}(\xi_{\tau_n})/(1-\tau_n))}{j!} [n(1-\tau_n)]^{(1-j)/2} \left\{ \sqrt{n(1-\tau_n)} \left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} - \frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n} \right) \right\}^j .
\end{aligned}$$

Evaluating the variance of the right-hand side and its covariance with $\sqrt{n(1-\tau_n)}(\widehat{\xi}_{\tau_n}/\xi_{\tau_n} - 1)$ and truncating the resulting (formal) power series at an order $1/[n(1-\tau_n)]^J$, for a suitably chosen $J \geq 1$, will lead to a matrix $\mathfrak{M}(J) = \mathfrak{M}_n(J)$ whose first two rows and columns will represent a correction for the asymptotic variance matrix in (B7), with higher values of J intuitively linked to a more accurate approximation. Of course, first of all, $\mathfrak{M}_{n,22}(J) = M_{n,22}$ since this asymptotic variance of $\sqrt{n(1-\tau_n)}(\widehat{\xi}_{\tau_n}/\xi_{\tau_n} - 1)$ does not involve the above power series. Then, recalling that

$$\sqrt{n(1-\tau_n)} \left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} - \frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n} \right) \approx \sqrt{M_{n,11}}\, Z$$

where Z is standard Gaussian, $\mathfrak{M}_{n,11}(J)$ is obtained by calculating first

$$\mathfrak{M}_{n,11}(\infty) = \mathrm{Var}\left(\sum_{j=1}^{\infty} \frac{\psi^{(j)}(\overline{F}(\xi_{\tau_n})/(1-\tau_n))}{j!} [n(1-\tau_n)]^{(1-j)/2} M^{j/2}_{n,11} Z^j \right).$$

Recall that $\mathbb{E}(Z^j) = 0$ if j is odd and $(j-1)!!$ if j is even, where the double factorial $N!!$ denotes the product of all integers from 1 to N having the same parity as N, *i.e.*, when N is even, $(N-1)!! = (N-1)(N-3)\cdots 3 \cdot 1 = N!/(2^{N/2}(N/2)!)$. It follows that

$$\mathrm{Cov}(Z^i, Z^j) = \begin{cases} 0 & \text{when } i+j \text{ is odd,} \\ (i+j-1)!! & \text{when } i \text{ and } j \text{ are odd,} \\ (i+j-1)!! - (i-1)!!(j-1)!! & \text{when } i \text{ and } j \text{ are even.} \end{cases}$$

Paired with the fact that $\psi^{(j)}(x) = (-1)^j j!(1+x)^{-j-1}$, we obtain, after straightforward calculations and gathering together negative powers of $n(1-\tau_n)$,

$$
\begin{aligned}
&\mathfrak{M}_{n,11}(\infty) = (1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-4} M_{n,11} \\
&\times \left(1 + \sum_{j=1}^{\infty} \frac{(1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-2j}}{[n(1-\tau_n)]^j} M_{n,11}^j \left[(2j+1)!!(2j+1) - \sum_{i=1}^{j}(2i-1)!!(2j+1-2i)!!\right]\right) \\
&=: \sum_{j=0}^{\infty} \frac{\mathfrak{m}_{j,11}}{[n(1-\tau_n)]^j}.
\end{aligned}
$$

We next seek $\mathfrak{M}_{n,12}(J)$, which requires evaluating

$$
\operatorname{Cov}\left(\left\{\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} - \frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right)\right\}^j, \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right), \quad \text{for } j \geq 1.
$$

Recall that if (X_1, X_2) is a Gaussian centered random pair then $\mathbb{E}(X_2|X_1) = (\operatorname{Cov}(X_1,X_2)/\operatorname{Var}(X_1))X_1$. It follows that, for any $j \geq 1$,

$$
\operatorname{Cov}(X_1^j, X_2) = \mathbb{E}(X_1^j X_2) = \frac{\operatorname{Cov}(X_1,X_2)}{\operatorname{Var}(X_1)} \mathbb{E}(X_1^{j+1}).
$$

Therefore

$$
\begin{aligned}
&\mathfrak{M}_{n,12}(\infty) \\
&= \sum_{j=1}^{\infty} \frac{\psi^{(j)}(\overline{F}(\xi_{\tau_n})/(1-\tau_n))}{j!} [n(1-\tau_n)]^{(1-j)/2} M_{n,12} M_{n,11}^{(j-1)/2} \mathbb{E}(Z^{j+1}) \\
&= -(1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-2} M_{n,12} \left(1 + \sum_{j=1}^{\infty} \frac{(1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-2j}}{[n(1-\tau_n)]^j} M_{n,11}^j (2j+1)!!\right) \\
&=: \sum_{j=0}^{\infty} \frac{\mathfrak{m}_{j,12}}{[n(1-\tau_n)]^j}.
\end{aligned}
$$

The matrix $\mathfrak{M}(J) = \mathfrak{M}_n(J)$, approximating the covariance matrix of $\sqrt{n(1-\tau_n)}(\widehat{\gamma}_{\tau_n}^{\mathrm{E}} - \psi(\overline{F}(\xi_{\tau_n})/(1-\tau_n)), \widehat{\xi}_{\tau_n}/\xi_{\tau_n} - 1)$, is finally obtained by truncating each of the series defining $\mathfrak{M}_{n,11}(\infty)$ and $\mathfrak{M}_{n,12}(\infty)$ at order $1/[n(1-\tau_n)]^J$. We now use this in order to approximate the uncertainty about $\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}$. One has

$$
\begin{aligned}
\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}} &= \left(1 + (2\tau_n - 1)(1/\widehat{\gamma}_{\tau_n}^{\mathrm{E}} - 1)\left(1 - \frac{\overline{Y}_n}{\widehat{\xi}_{\tau_n}}\right)^{-1}(1 + \mathrm{O}_{\mathbb{P}}(A((1-\tau_n)^{-1})))\right)^{-1} \\
&= u_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E}}, \widehat{\xi}_{\tau_n}/\xi_{\tau_n}) + \mathrm{O}_{\mathbb{P}}(A((1-\tau_n)^{-1})),
\end{aligned}
$$

where

$$
u_n(x,y) = \left(1 + (2\tau_n - 1)\left(\frac{1}{x} - 1\right)\left(1 - \frac{\mathbb{E}(Y)}{y\xi_{\tau_n}}\right)^{-1}\right)^{-1}.
$$

Here and in the rest of this section, we have neglected the finite-sample variability in $\overline{Y}_n$, which typically converges faster to $\mathbb{E}(Y)$ than the other (extreme value) terms do to their respective limits. We have also neglected any term proportional to $A((1-\tau_n)^{-1})$ or $A((1-\tau_n')^{-1})$. The rationale behind this choice is that, since $A(t) = b\gamma t^{\rho}$, keeping these terms in this kind of calculation would then entail approximating the correlation of estimators of the second-order parameters ρ and b with estimators of other extreme value parameters, here $\widehat{\gamma}_{\tau_n}^{\mathrm{E}}$ and $\widehat{\xi}_{\tau_n}$. This is a joint convergence problem which, to the best of our knowledge, remains open, and whose solution deserves a separate in-depth study. We now write

$$
\begin{aligned}
&\sqrt{n(1-\tau_n)}\left(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}-\gamma, \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right) \\
&= \sqrt{n(1-\tau_n)}\left(u_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E}}, \widehat{\xi}_{\tau_n}/\xi_{\tau_n}) - u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1}, 1), \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right) + \mathrm{o}_{\mathbb{P}}(1).
\end{aligned}
$$

The random vector

$$
\sqrt{n(1-\tau_n)}\left(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}-\gamma, \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)
$$

can thus be considered asymptotically Gaussian centered with a covariance matrix approximated by $\mathfrak{M}^{\mathrm{BR}}(J) = \mathfrak{M}_n^{\mathrm{BR}}(J)$ whose elements are

$$
\begin{aligned}
\mathfrak{M}_{n,11}^{\mathrm{BR}}(J) &= (\partial_1 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1}, 1))^2 \mathfrak{M}_{n,11}(J) \\
&\quad + (\partial_2 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1}, 1))^2 \mathfrak{M}_{n,22}(J) \\
&\quad + 2\partial_1 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1}, 1)\partial_2 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1}, 1)\mathfrak{M}_{n,12}(J), \\
\mathfrak{M}_{n,12}^{\mathrm{BR}}(J) &= \partial_1 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1}, 1)\mathfrak{M}_{n,12}(J) \\
&\quad + \partial_2 u_n((1+\overline{F}(\xi_{\tau_n})/(1-\tau_n))^{-1}, 1)\mathfrak{M}_{n,22}(J), \\
\mathfrak{M}_{n,22}^{\mathrm{BR}}(J) &= \mathfrak{M}_{n,22}(J).
\end{aligned}
$$

Here

$$
\partial_1 u_n(x,y) = \frac{\xi_{\tau_n}(2\tau_n-1)(\xi_{\tau_n}y - \mathbb{E}(Y))y}{[(\mathbb{E}(Y)-2\xi_{\tau_n}y(1-\tau_n))x - \xi_{\tau_n}y(2\tau_n-1)]^2} \tag{B8}
$$

$$
\text{and } \partial_2 u_n(x,y) = \frac{\xi_{\tau_n}(2\tau_n-1)\mathbb{E}(Y)x(1-x)}{[(\mathbb{E}(Y)-2\xi_{\tau_n}y(1-\tau_n))x - \xi_{\tau_n}y(2\tau_n-1)]^2}. \tag{B9}
$$

Our final step is to combine all these elements in order to accurately quantify the uncertainty in $\log(\widehat{\xi}_{\tau_n'}^{\star,\mathrm{BR}}/\xi_{\tau_n'})$. We rewrite this quantity as

$$
\begin{aligned}
\log\frac{\widehat{\xi}_{\tau_n'}^{\star,\mathrm{BR}}}{\xi_{\tau_n'}} &= (\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}-\gamma)\log\left(\frac{1-\tau_n}{1-\tau_n'}\right) + \log\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - \log(1+B_{1,n}) \\
&\quad + \widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}\log\left(\frac{1-\overline{Y}_n/\widehat{\xi}_{\tau_n}}{2\tau_n-1}\right) - \log(1+B_{2,n}) \\
&\quad - \widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}\log\left(\frac{1-\overline{Y}_n/\left(\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}}\widehat{\xi}_{\tau_n}\right)}{2\tau_n'-1}\right) - \log(1+B_{3,n}) \\
&\quad + \mathrm{O}_{\mathbb{P}}(A((1-\tau_n)^{-1})).
\end{aligned}
$$

Since we neglect any term proportional to $A((1-\tau_n)^{-1})$ or $A((1-\tau_n')^{-1})$, we now:

- Neglect $B_{1,n}$,
- Approximate $1+B_{2,n}$ by $(1+r(\tau_n))^{\gamma}$, where $1+r(\tau_n)$ is itself approximated by $(1-\mathbb{E}(Y)/\xi_{\tau_n})/(2\tau_n-1)$,
- Approximate $1+B_{3,n}$ by $(1+r(\tau_n'))^{-\gamma}$, where $1+r(\tau_n')$ is itself approximated by $(1-\mathbb{E}(Y)/\xi_{\tau_n'})/(2\tau_n'-1)$.

Hence the approximation

$$
\begin{aligned}
\log \frac{\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\xi_{\tau_n'}} &\approx \left[\log\left(\frac{1-\tau_n}{1-\tau_n'}\right) + \log\left(\frac{2\tau_n'-1}{2\tau_n-1}\right)\right] (\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n} - \gamma) + \log \frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} \\
&+ \widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n} \log\left(1 - \frac{\mathbb{E}(Y)}{\widehat{\xi}_{\tau_n}}\right) - \gamma \log\left(1 - \frac{\mathbb{E}(Y)}{\xi_{\tau_n}}\right) \\
&- \left[\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n} \log\left(1 - \frac{\mathbb{E}(Y)}{\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}} \widehat{\xi}_{\tau_n}}\right) - \gamma \log\left(1 - \frac{\mathbb{E}(Y)}{\xi_{\tau_n'}}\right)\right].
\end{aligned}
$$

We finally write

$$
\begin{aligned}
&\frac{1}{\xi_{\tau_n'}} - \frac{1}{\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma} \xi_{\tau_n}} \\
&= \frac{1}{\xi_{\tau_n'}} \left(1 - \frac{\xi_{\tau_n'}}{q_{\tau_n'}} \times \frac{q_{\tau_n}}{\xi_{\tau_n}} \times \frac{q_{\tau_n'}}{\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma} q_{\tau_n}}\right) \\
&= \mathrm{O}\left(\frac{1}{\xi_{\tau_n'}} \left(\mathrm{O}(|A((1-\tau_n)^{-1})|) + \mathrm{O}(|A((1-\tau_n')^{-1})|) + \frac{1}{\xi_{\tau_n}} + \frac{1}{\xi_{\tau_n'}}\right)\right)
\end{aligned}
$$

which is typically very small due to the presence of the factor $1/\xi_{\tau_n'}$ (the inverse of an expectile at the properly extreme level τ_n'). This motivates approximating $\log(1-\mathbb{E}(Y)/\xi_{\tau_n'})$ by

$$
\log\left(1 - \frac{\mathbb{E}(Y)}{\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma} \xi_{\tau_n}}\right).
$$

This eventually leads to the approximation

$$
\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))} \log \frac{\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\xi_{\tau_n'}} \approx \sqrt{n(1-\tau_n)} \left\{ g_n(\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}, \widehat{\xi}_{\tau_n}/\xi_{\tau_n}) - g_n(\gamma, 1) \right\}
$$

where

$$
\begin{aligned}
g_n(x,y) &= x\left(1 + \frac{\log((2\tau_n'-1)/(2\tau_n-1))}{\log((1-\tau_n)/(1-\tau_n'))}\right) + \frac{\log(y)}{\log((1-\tau_n)/(1-\tau_n'))} \\
&+ \frac{x}{\log((1-\tau_n)/(1-\tau_n'))} \log\left(1 - \frac{\mathbb{E}(Y)}{\xi_{\tau_n} y}\right)
\end{aligned}
$$

$$- \frac{x}{\log((1-\tau_n)/(1-\tau_n'))} \log\left(1 - \frac{\mathbb{E}(Y)}{\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-x} \xi_{\tau_n} y}\right).$$

The variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ is then well approximated by

$$s_n^2(J) = (\partial_1 g_n(\gamma,1))^2 \mathfrak{M}^{\mathrm{BR}}_{n,11}(J) + 2\partial_1 g_n(\gamma,1)\partial_2 g_n(\gamma,1)\mathfrak{M}^{\mathrm{BR}}_{n,12}(J) + (\partial_2 g_n(\gamma,1))^2 \mathfrak{M}^{\mathrm{BR}}_{n,22}(J)$$

where

$$\begin{aligned} &\partial_1 g_n(\gamma,1) \\ &= 1 + \frac{\log((2\tau_n'-1)/(2\tau_n-1))}{\log((1-\tau_n)/(1-\tau_n'))} + \frac{\log(1-\mathbb{E}(Y)/\xi_{\tau_n})}{\log((1-\tau_n)/(1-\tau_n'))} \\ &- \frac{\log(1-\mathbb{E}(Y)/(((1-\tau_n')/(1-\tau_n))^{-\gamma}\,\xi_{\tau_n}))}{\log((1-\tau_n)/(1-\tau_n'))} - \frac{\gamma\mathbb{E}(Y)}{((1-\tau_n')/(1-\tau_n))^{-\gamma}\,\xi_{\tau_n} - \mathbb{E}(Y)} \end{aligned} \tag{B10}$$

and

$$\begin{aligned} &\partial_2 g_n(\gamma,1) \\ &= \frac{1}{\log((1-\tau_n)/(1-\tau_n'))}\left(1 - \frac{\gamma\mathbb{E}(Y)}{((1-\tau_n')/(1-\tau_n))^{-\gamma}\,\xi_{\tau_n} - \mathbb{E}(Y)} + \frac{\gamma\mathbb{E}(Y)}{\xi_{\tau_n} - \mathbb{E}(Y)}\right). \end{aligned} \tag{B11}$$

In order to produce confidence intervals, we estimate γ by $\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}$, ξ_{τ_n} by $\widehat{\xi}_{\tau_n}$, $\mathbb{E}(Y)$ by $\overline{Y}_n$, $\overline{F}(\xi_{\tau_n})$ by $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$, $\varphi^{(1)}(\xi_{\tau_n})$ by

$$\widehat{\varphi}^{(1)}_n(\widehat{\xi}_{\tau_n}) = \frac{1}{n}\sum_{i=1}^n (Y_i - \widehat{\xi}_{\tau_n})\mathbb{1}\{Y_i > \widehat{\xi}_{\tau_n}\}.$$

The estimation of $\varphi^{(2)}(\xi_{\tau_n})$ is more complex, because its naive empirical counterpart $\widehat{\varphi}^{(2)}_n(\widehat{\xi}_{\tau_n})$ is unbiased but highly skewed, and therefore tends to vastly underestimate $\varphi^{(2)}(\xi_{\tau_n})$. A second-order approximation of the underlying distribution function $\overline{F}$ in a neighborhood of infinity suggests the expansion

$$\varphi^{(2)}(\xi_{\tau_n}) \approx 2\overline{F}(\xi_{\tau_n})\xi_{\tau_n}^2\gamma^2\left(\frac{1}{(1-\gamma)(1-2\gamma)} + \frac{b\overline{F}(\xi_{\tau_n})^{-\rho}}{\rho}\left\{\frac{1}{(1-\gamma-\rho)(1-2\gamma-\rho)} - \frac{1}{(1-\gamma)(1-2\gamma)}\right\}\right).$$

A simple plug-in then suggests the estimator

$$\widetilde{\varphi}^{(2)}_n(\widehat{\xi}_{\tau_n}) = 2\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})\widehat{\xi}^2_{\tau_n}(\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})^2\left(\frac{1}{(1-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})(1-2\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})} + \frac{\overline{b}(\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n}))^{-\overline{\rho}}}{\overline{\rho}}\left\{\frac{1}{(1-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}-\overline{\rho})(1-2\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}-\overline{\rho})} - \frac{1}{(1-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})(1-2\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})}\right\}\right)$$

which tends to be much more stable in finite samples than $\widehat{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n})$. We deduce from these calculations a corrected asymptotic Gaussian confidence interval for $\xi_{\tau_n'}$ at level $1-\alpha$ as

$$\widehat{I}_{\tau_n'}^{(2)}(\alpha) = \widehat{I}_{\tau_n'}^{(2)}(\alpha; J) = \left[\widehat{\xi}_{\tau_n'}^{\star,\mathrm{BR}} \exp\left(\pm \frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}} \sqrt{\widehat{s}_n^2(J)} \times z_{1-\alpha/2} \right) \right]$$

where $\widehat{s}_n^2(J)$ is obtained from $s_n^2(J)$ by replacing γ by $\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}$, ξ_{τ_n} by $\widehat{\xi}_{\tau_n}$, $\mathbb{E}(Y)$ by $\overline{Y}_n$, $\overline{F}(\xi_{\tau_n})$ by $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$, $\varphi^{(1)}(\xi_{\tau_n})$ by $\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n})$ and $\varphi^{(2)}(\xi_{\tau_n})$ by $\widetilde{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n})$.

B.2 Quantile-based confidence interval construction

B.2.1 Preliminary steps

The quantile-based intermediate expectile estimator is $\widetilde{\xi}_{\tau_n} = (1/\widehat{\gamma}_{\tau_n}^{\mathrm{H}} - 1)^{-\widehat{\gamma}_{\tau_n}^{\mathrm{H}}} \widehat{q}_{\tau_n}$. Here we take $\tau_n = 1 - k_n/n$, where k_n is a sequence of integers, $\widehat{q}_{\tau_n} = Y_{n-\lfloor n(1-\tau_n)\rfloor,n} = Y_{n-k_n,n}$ is the corresponding intermediate order statistic, and

$$\widehat{\gamma}_{\tau_n}^{\mathrm{H}} = \widehat{\gamma}_{1-k_n/n}^{\mathrm{H}} = \frac{1}{k_n} \sum_{i=1}^{k_n} \log Y_{n-i+1,n} - \log Y_{n-k_n,n}$$

is the usual Hill estimator of γ calculated upon the top k_n log-spacings. Then

$$\log\left(\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}} \right) = \phi(\widehat{\gamma}_{1-k_n/n}^{\mathrm{H}}) - \phi(\gamma) + \log\left(\frac{\widehat{q}_{1-k_n/n}}{q_{1-k_n/n}} \right). \tag{B12}$$

where $\phi(x) = -x\log(x^{-1}-1)$. It is known that, if $\lambda_1 = \lim_{n\to\infty} \sqrt{k_n} A(n/k_n) \in \mathbb{R}$,

$$\sqrt{k_n}\left(\widehat{\gamma}_{1-k_n/n}^{\mathrm{H}} - \gamma, \log\left(\frac{\widehat{q}_{1-k_n/n}}{q_{1-k_n/n}} \right) \right) \xrightarrow{d} \mathcal{N}\left(\left(\frac{\lambda_1}{1-\rho}, 0 \right), \gamma^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

(combine *e.g.* Theorem 2.4.8 p.52, Lemma 3.2.3 p.71 and Theorem 3.2.5 p.74 in de Haan and Ferreira (2006)). The asymptotic correlation of the second term in (B12) with the first one is then exactly 0, and the third term is a bias term that is evaluated by employing Proposition 1 in Daouia et al. (2020). A straightforward application of the delta-method with the function $g(x,y) = (x, y - x\log(x^{-1}-1))$ leads to:

$$\sqrt{k_n}\left(\widehat{\gamma}_{1-k_n/n}^{\mathrm{H}} - \gamma, \log\left(\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}} \right) \right) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda_1}{1-\rho}(1, m(\gamma)), \mathbf{V} \right), \tag{B13}$$

where $\lambda_1 = \lim_{n\to\infty} \sqrt{k_n} A(n/k_n)$, $\lambda_2 = \lim_{n\to\infty} \sqrt{k_n}/q_{1-k_n/n}$,

$$B_1 = \frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} + \frac{(\gamma^{-1}-1)^{-\rho} - 1}{\rho} \text{ and } B_2 = \gamma(\gamma^{-1}-1)^{\gamma}\mathbb{E}(Y),$$

with $m(\gamma) = \phi'(\gamma) = (1-\gamma)^{-1} - \log(\gamma^{-1}-1)$, and $\mathbf{V}$ is the symmetric matrix having entries $V_{11} = \gamma^2$, $V_{12} = \gamma^2 m(\gamma)$ and $V_{22} = \gamma^2(1 + (m(\gamma))^2)$. Write finally

$$\log\left(\frac{\widetilde{\xi}_{\tau_n'}^{\star}}{\xi_{\tau_n'}} \right) = (\widehat{\gamma}_{1-k_n/n}^{\mathrm{H}} - \gamma) \log\left(\frac{k_n}{n(1-\tau_n')} \right) + \log\left(\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}} \right)$$

$$+\log\left((\gamma^{-1}-1)^{-\gamma}\frac{q_{\tau_n'}}{\xi_{\tau_n'}}\right)+\log\left(\left[\frac{n(1-\tau_n')}{k_n}\right]^{-\gamma}\frac{q_{1-k_n/n}}{q_{\tau_n'}}\right). \tag{B14}$$

Convergence (2.4) follows using Proposition 1 in Daouia et al. (2020) (for the first nonrandom term in the second line of (B14)) and the equation at the top of p.139 in the proof of Theorem 4.3.8 in de Haan and Ferreira (2006) (for the second nonrandom term), and since $\overline{b}-b=\mathrm{o}_{\mathbb{P}}(1)$ and $(\overline{\rho}-\rho)\log(n)=\mathrm{o}_{\mathbb{P}}(1)$, $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ satisfies the same asymptotic normality result but with asymptotic mean 0.

B.2.2 Detailed calculation

Convergence (2.4) is obtained by neglecting the finite-sample uncertainty in the estimator $\widetilde{\xi}_{1-k_n/n}$. Combining (B13) and (B14) suggests that the asymptotic variance of

$$\frac{\sqrt{k_n}}{\log\left(k_n/(n(1-\tau_n'))\right)}\log\left(\frac{\widetilde{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}}\right)$$

may be better approximated by

$$\begin{aligned}V_{11}+2\frac{V_{12}}{\log\left(k_n/(n(1-\tau_n'))\right)}+\frac{V_{22}}{\log^2\left(k_n/(n(1-\tau_n'))\right)}&\\ =\gamma^2\left[1+2\frac{m(\gamma)}{\log\left(k_n/(n(1-\tau_n'))\right)}+\frac{1+(m(\gamma))^2}{\log^2\left(k_n/(n(1-\tau_n'))\right)}\right].&\end{aligned} \tag{B15}$$

This approximation substantially improves upon the expression of the variance obtained through first-order asymptotics, but behaves somewhat poorly when γ is in a neighborhood of 1. An even more accurate approximation is possible by using the same trick as in Section B.1 and writing a formal power series expansion of ϕ instead. Use (B12) in order to write

$$\begin{aligned}&\sqrt{k_n}\log\left(\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n}}\right)\\ &\qquad=\sqrt{k_n}\log\left(\frac{\widehat{q}_{1-k_n/n}}{q_{1-k_n/n}}\right)+\sum_{j=1}^{\infty}\frac{\phi^{(j)}(\gamma)}{j!}k_n^{\frac{1-j}{2}}\left\{\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma)\right\}^j.\end{aligned}$$

As in Section B.1, we approximate the variance of $\sqrt{k_n}\log(\widetilde{\xi}_{1-k_n/n}/((\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n}))$ and its covariance with $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma)$ by working with the above right-hand side in a formal way and truncating the resulting (formal) variance and covariance power series at an order $1/k_n^J$, for a suitably chosen $J\geq 1$. This provides a finite-sample correction $\mathbf{V}(J)=\mathbf{V}_n(J)$ for the matrix $\mathbf{V}=\mathbf{V}(0)$ in (B13), and hence a further correction for the asymptotic variance of the extrapolated quantile-based extreme expectile estimator. The resulting matrix $\mathbf{V}(J)$ has the following coefficients: the first variance term $V_{11}(J)=V_{11}(0)=\gamma^2$, obtained from the asymptotic variance of $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma)$, does not depend on J. To calculate the second variance term $V_{22}(J)$, obtained by evaluating the variance of the right-hand side in the above expansion and Equation (B13), recall first that $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}-\gamma,\widehat{q}_{1-k_n/n}/q_{1-k_n/n}-1)$ converges to a Gaussian random pair with independent components having variance γ^2. We shall then obtain $V_{22}(J)$ after the calculation of

$$V_{22}(\infty)=\gamma^2+\operatorname{Var}\left(\sum_{j=1}^{\infty}\frac{\phi^{(j)}(\gamma)}{j!}k_n^{\frac{1-j}{2}}\gamma^jZ^j\right)$$

where Z is standard Gaussian. One finds

$$
\begin{aligned}
V_{22}(\infty) &= \gamma^2 + \sum_{\substack{i,j\geq 1 \\ i+j \text{ even}}} \frac{\phi^{(i)}(\gamma)\phi^{(j)}(\gamma)(i+j-1)!!}{i!j!\,k_n^{(i+j)/2-1}}\gamma^{i+j} \\
&- \sum_{\substack{i,j\geq 1 \\ i \text{ and } j \text{ even}}} \frac{\phi^{(i)}(\gamma)\phi^{(j)}(\gamma)}{i!j!\,k_n^{(i+j)/2-1}}(i-1)!!(j-1)!!\gamma^{i+j} \\
&= \gamma^2(1+(m(\gamma))^2) + \sum_{\substack{i,j\geq 1 \\ i+j>2 \\ i+j \text{ even}}} \frac{\phi^{(i)}(\gamma)\phi^{(j)}(\gamma)(i+j-1)!!}{i!j!\,k_n^{(i+j)/2-1}}\gamma^{i+j} \\
&- \sum_{i,j\geq 1} \frac{\phi^{(2i)}(\gamma)\phi^{(2j)}(\gamma)}{i!j!2^{i+j}k_n^{i+j-1}}\gamma^{2(i+j)}.
\end{aligned}
$$

Gathering powers of k_n together, we find

$$
\begin{aligned}
V_{22}(\infty) &= \gamma^2(1+(m(\gamma))^2) + \sum_{j=1}^{\infty} \frac{(2j+1)!!\gamma^{2j+2}}{k_n^j} \sum_{i=1}^{2j+1} \frac{\phi^{(i)}(\gamma)\phi^{(2j+2-i)}(\gamma)}{i!(2j+2-i)!} \\
&- \sum_{j=1}^{\infty} \frac{\gamma^{2j+2}}{2^{j+1}k_n^j} \sum_{i=1}^{j} \frac{\phi^{(2i)}(\gamma)\phi^{(2j+2-2i)}(\gamma)}{i!(j+1-i)!} \\
&= \gamma^2 \left(1+(m(\gamma))^2 + \sum_{j=1}^{\infty} \frac{\gamma^{2j}}{k_n^j} \left((2j+1)!! \sum_{i=1}^{2j+1} \frac{\phi^{(i)}(\gamma)\phi^{(2j+2-i)}(\gamma)}{i!(2j+2-i)!} \right.\right. \\
&\left.\left. - \frac{1}{2^{j+1}} \sum_{i=1}^{j} \frac{\phi^{(2i)}(\gamma)\phi^{(2j+2-2i)}(\gamma)}{i!(j+1-i)!}\right)\right) =: \sum_{j=0}^{\infty} \frac{v_{j,22}}{k_n^j}.
\end{aligned}
$$

An expression of the formal covariance term $V_{12}(\infty)$ is similarly easily derived: if Z is standard Gaussian,

$$
\begin{aligned}
V_{12}(\infty) = \sum_{j=0}^{\infty} \frac{\phi^{(j)}(\gamma)}{j!} k_n^{\frac{1-j}{2}} \gamma^{j+1}\mathbb{E}(Z^{j+1}) &= \sum_{\substack{j\geq 0 \\ j \text{ odd}}} \frac{\phi^{(j)}(\gamma)}{j!} k_n^{\frac{1-j}{2}} \gamma^{j+1} j!! \\
&= \sum_{j=0}^{\infty} \frac{\phi^{(2j+1)}(\gamma)}{2^j j! k_n^j}\gamma^{2j+2} =: \sum_{j=0}^{\infty} \frac{v_{j,12}}{k_n^j}.
\end{aligned}
$$

Closed-form expressions of the coefficients $v_{j,12}$ and $v_{j,22}$, for $j \geq 1$, can then be obtained by remarking that

$$
\phi''(\gamma) = m'(\gamma) = \frac{1}{(1-\gamma)^2} + \frac{1}{\gamma(1-\gamma)} = \frac{1}{\gamma} + \frac{1}{1-\gamma} + \frac{1}{(1-\gamma)^2}
$$

and therefore

$$
\begin{aligned}
\forall j \geq 2,\ \phi^{(j)}(\gamma) &= \frac{(-1)^{j-2}(j-2)!}{\gamma^{j-1}} + \frac{(j-2)!}{(1-\gamma)^{j-1}} + \frac{(j-1)!}{(1-\gamma)^j} \\
&= \frac{(-1)^{j-2}(j-2)! \sum_{i=0}^{j-2} \binom{j}{i}(-1)^i\gamma^i}{\gamma^{j-1}(1-\gamma)^j}.
\end{aligned}
$$

In practice we truncate the power series at order $1/k_n^J$, for a given $J \geq 1$, so that a refined approximation to the covariance matrix of $\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n} - \gamma, \log(\widetilde{\xi}_{1-k_n/n}/((\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}))$ is the symmetric matrix $\mathbf{V}(J)$, where the elements of $\mathbf{V}(J)$ are

$$V_{11}(J) = \gamma^2, \ V_{12}(J) = \sum_{j=0}^{J} \frac{v_{j,12}}{k_n^j}, \ V_{22}(J) = \sum_{j=0}^{J} \frac{v_{j,22}}{k_n^j}.$$

In our implementation we stop the approximation at order $1/k_n^2$, that is, $J = 2$. A use of this refinement will be shown to dramatically improve the coverage probabilities of the asymptotic Gaussian confidence interval, especially in the challenging case when γ is close to 1 and the right tail of Y is very heavy. The explicit expressions of the coefficients $v_{j,12}$ and $v_{j,22}$, for $j = 0, 1, 2$, are provided in Table B1.

j	$v_{j,12}/\gamma^2$	$v_{j,22}/\gamma^2$
0	$m(\gamma)$	$(1 + (m(\gamma))^2)$
1	$\dfrac{3\gamma - 1}{2(1-\gamma)^3}$	$\dfrac{(3\gamma-1)m(\gamma)}{(1-\gamma)^3} + \dfrac{1}{2(1-\gamma)^4}$
2	$\dfrac{3(10\gamma^3 - 10\gamma^2 + 5\gamma - 1)}{4(1-\gamma)^5}$	$\dfrac{3(10\gamma^3 - 10\gamma^2 + 5\gamma - 1)m(\gamma)}{2(1-\gamma)^5} + \dfrac{6\gamma^2 - 4\gamma + 1}{(1-\gamma)^6} + \dfrac{5(3\gamma-1)^2}{12(1-\gamma)^6}$

Table B1: First values of the coefficients $v_{j,12}/\gamma^2$ and $v_{j,22}/\gamma^2$, for $j = 0, 1, 2$.

We may now proceed with the construction of a confidence interval about $\xi_{\tau'_n}$. Recall Equation (2.6), which can be rewritten as

$$\log \frac{\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}}{\xi_{\tau'_n}} = \log\left(\frac{k_n}{n(1-\tau'_n)}\right)(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n} - \gamma) + \log\left(\frac{(1/\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n} - 1)^{-\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n}} \widehat{q}_{1-k_n/n}}{(1/\gamma - 1)^{-\gamma} q_{1-k_n/n}}\right)$$
$$+ \log\left((\gamma^{-1}-1)^{-\gamma}\frac{q_{\tau'_n}}{\xi_{\tau'_n}}\right) + \log\left(\left[\frac{n(1-\tau'_n)}{k_n}\right]^{-\gamma} \frac{q_{1-k_n/n}}{q_{\tau'_n}}(1+\widetilde{B}_{1,n})(1+\widetilde{B}_{3,n})\right).$$

Carrying out accurate inference about $\xi_{\tau'_n}$ based on $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}$ requires providing a good approximation of the variance of the right-hand side. For that purpose, as in Section B.1, we neglect any term proportional to $A(n/k_n)$ or $A((1-\tau'_n)^{-1})$. We therefore:

- Approximate $\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n}$ by $\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}$,
- Write
$$\log\left((\gamma^{-1}-1)^{-\gamma}\frac{q_{\tau'_n}}{\xi_{\tau'_n}}\right) \approx \gamma\log\left(1 - \frac{\mathbb{E}(Y)}{\xi_{\tau'_n}}\right) - \gamma\log(2\tau'_n - 1),$$
see Equation (12) in Girard et al. (2022),
- Neglect $\log([n(1-\tau'_n)/k_n]^{-\gamma} q_{1-k_n/n}/q_{\tau'_n})$, in virtue of the equation at the top of p.139 in de Haan and Ferreira (2006),
- Neglect $\widetilde{B}_{1,n}$,
- Approximate $1 + \widetilde{B}_{3,n}$ by $(1 + \widetilde{r}^{\star}(\tau'_n))^{-\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}}$, where $1 + \widetilde{r}^{\star}(\tau'_n)$ is itself approximated by $(1 - \mathbb{E}(Y)/\widetilde{\xi}^{\star}_{\tau'_n})/(2\tau'_n - 1)$.

This suggests the simpler asymptotic approximation

$$\log \frac{\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}}{\xi_{\tau'_n}} \approx \left(\log\left(\frac{k_n}{n(1-\tau'_n)}\right) + \log(2\tau'_n - 1)\right)(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n} - \gamma) + \log\frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}}$$

$$- \left(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n} \log\left(1 - \frac{\mathbb{E}(Y)}{\widetilde{\xi}^{\star}_{\tau'_n}}\right) - \gamma \log\left(1 - \frac{\mathbb{E}(Y)}{\xi_{\tau'_n}}\right)\right).$$

We finally write

$$\begin{aligned}
&\frac{1}{\xi_{\tau'_n}} - \frac{1}{\left(\frac{n(1-\tau'_n)}{k_n}\right)^{-\gamma}(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}} \\
&= \frac{1}{\xi_{\tau'_n}}\left(1 - (\gamma^{-1}-1)^{\gamma}\frac{\xi_{\tau'_n}}{q_{\tau'_n}} \times \frac{q_{\tau'_n}}{\left(\frac{n(1-\tau'_n)}{k_n}\right)^{-\gamma} q_{1-k_n/n}}\right) \\
&= \mathrm{O}\left(\frac{1}{\xi_{\tau'_n}}\left(\mathrm{O}(|A(n/k_n)|) + \mathrm{O}(|A((1-\tau'_n)^{-1})|) + \frac{1}{\xi_{\tau'_n}}\right)\right)
\end{aligned}$$

which is typically very small due to the presence of the factor $1/\xi_{\tau'_n}$ (the inverse of an expectile at the properly extreme level τ'_n). This motivates approximating $\log(1-\mathbb{E}(Y)/\xi_{\tau'_n})$ by

$$\log\left(1 - \frac{\mathbb{E}(Y)}{\left(\frac{n(1-\tau'_n)}{k_n}\right)^{-\gamma}(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}}\right).$$

This eventually leads to the approximation

$$\frac{\sqrt{k_n}}{\log\left(k_n/(n(1-\tau'_n))\right)} \log\frac{\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}}{\xi_{\tau'_n}} \approx \sqrt{k_n}\left\{h_n\left(\widehat{\gamma}^{\mathrm{H}}_{1-k_n/n}, \frac{\widetilde{\xi}_{1-k_n/n}}{(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}}\right) - h_n(\gamma, 1)\right\}$$

where

$$\begin{aligned}
h_n(x,y) &= \left(1 + \frac{\log(2\tau'_n - 1)}{\log(k_n/(n(1-\tau'_n)))}\right) x + \frac{\log(y)}{\log(k_n/(n(1-\tau'_n)))} \\
&\quad - \frac{x}{\log(k_n/(n(1-\tau'_n)))} \log\left(1 - \frac{\mathbb{E}(Y)}{\left(\frac{n(1-\tau'_n)}{k_n}\right)^{-x}(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}\, y}\right).
\end{aligned}$$

This suggests that the variance of $(\sqrt{k_n}/\log(k_n/(n(1-\tau'_n))))\log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}/\xi_{\tau'_n})$ is well approximated by

$$\sigma_n^2(J) = (\partial_1 h_n(\gamma,1))^2 V_{11}(J) + 2\partial_1 h_n(\gamma,1)\partial_2 h_n(\gamma,1) V_{12}(J) + (\partial_2 h_n(\gamma,1))^2 V_{22}(J)$$

where

$$\begin{aligned}
\partial_1 h_n(\gamma,1) &= 1 - \frac{\gamma\mathbb{E}(Y)}{\left(\frac{n(1-\tau'_n)}{k_n}\right)^{-\gamma}(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n} - \mathbb{E}(Y)} \\
&+ \frac{1}{\log(k_n/(n(1-\tau'_n)))}\left(\log(2\tau'_n - 1) - \log\left(1 - \frac{\mathbb{E}(Y)}{\left(\frac{n(1-\tau'_n)}{k_n}\right)^{-\gamma}(\gamma^{-1}-1)^{-\gamma} q_{1-k_n/n}}\right)\right) \qquad \text{(B16)}
\end{aligned}$$

and

$$\partial_2 h_n(\gamma,1) = \frac{1}{\log(k_n/(n(1-\tau_n')))}\left(1 - \frac{\gamma\mathbb{E}(Y)}{\left(\frac{n(1-\tau_n')}{k_n}\right)^{-\gamma}(\gamma^{-1}-1)^{-\gamma}q_{1-k_n/n} - \mathbb{E}(Y)}\right). \tag{B17}$$

Note that this is indeed a further refinement of (B15), since $\partial_1 h_n(\gamma,1)$ converges to 1 and $\partial_2 h_n(\gamma,1)$ is asymptotically equivalent to $1/\log(k_n/(n(1-\tau_n'))) \to 0$. We deduce from these calculations a corrected asymptotic Gaussian confidence interval for $\xi_{\tau_n'}$ at level $1-\alpha$ as

$$\widetilde{I}^{(2)}_{\tau_n'}(\alpha) = \widetilde{I}^{(2)}_{\tau_n'}(\alpha; J) = \left[\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'} \exp\left(\pm\frac{\log(k_n/(n(1-\tau_n')))}{\sqrt{k_n}}\sqrt{\widetilde{\sigma}_n^2(J)} \times z_{1-\alpha/2}\right)\right]$$

where

$$\widetilde{\sigma}_n^2(J) = (\partial_1\widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1))^2(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n})^2 + (\partial_2\widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1))^2\widehat{V}_{22}(J) \\ + 2\partial_1\widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)\partial_2\widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)\widehat{V}_{12}(J)$$

is obtained by replacing γ by $\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n}$, $q_{1-k_n/n}$ by $\widehat{q}_{1-k_n/n}$ and $\mathbb{E}(Y)$ by $\overline{Y}_n$ in the quantities $V_{11}(J) \equiv \gamma^2$, $V_{12}(J)$, $V_{22}(J)$, $\partial_1 h_n(\gamma,1)$ and $\partial_2 h_n(\gamma,1)$, thus producing their respective estimators $\widehat{V}_{11}(J) \equiv (\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n})^2$, $\widehat{V}_{12}(J)$, $\widehat{V}_{22}(J)$, $\partial_1\widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)$ and $\partial_2\widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)$.

Appendix C Theoretical results and proofs

C.1 Auxiliary results

All distributions in Table C2 have natural location parameter 0. Some of them, such as the Burr and Dagum distributions, have scale parameters which never appear in the values of ρ and b because of the invariance of the ratio of two quantiles by changes of scale. Location parameters can be handled using the following lemma.

Lemma 1. *Assume that a given distribution has a survival function $\overline{F}$ and a probability density function f (with respect to the Lebesgue measure) that satisfies*

$$f(t) = ct^{-\alpha-1}(1 + dt^{-\beta}(1+\mathrm{o}(1))) \text{ as } t \to \infty$$

where $\alpha, \beta, c > 0$ and $d \neq 0$. Set $\gamma = 1/\alpha$.
(i) Then $\overline{F}$ satisfies condition $\mathcal{C}_2(\gamma, -\beta/\alpha, A)$, where

$$A(t) = \left(-\frac{d\beta}{\alpha+\beta}(c/\alpha)^{-\beta/\alpha}\right) \times \frac{1}{\alpha} \times t^{-\beta/\alpha} = b\gamma t^{-\beta/\alpha}$$
$$\text{with } b = -\frac{d\beta}{\alpha+\beta}(c/\alpha)^{-\beta/\alpha}.$$

Assume moreover that

$$f(t) = ct^{-\alpha-1}(1 + dt^{-\beta} + d't^{-\beta-\beta'}(1+\mathrm{o}(1))) \text{ as } t \to \infty$$

where $\alpha, \beta, \beta', c > 0$ and $d, d' \neq 0$. Let $m \neq 0$ and define a survival function $\overline{G}$ having probability density function $g : t \mapsto f(t-m)$. Then:

(ii) If $\beta < 1$, then $\overline{G}$ satisfies condition $\mathcal{C}_2(\gamma, -\beta/\alpha, A)$, where A is as in (i),
(iii) If $\beta > 1$, then $\overline{G}$ satisfies condition $\mathcal{C}_2(\gamma, -1/\alpha, B_1)$, where

$$B_1(t) = -m(c/\alpha)^{-1/\alpha} \times \frac{1}{\alpha} \times t^{-1/\alpha} = b_1\gamma t^{-1/\alpha}$$
$$\textit{with } b_1 = -m(c/\alpha)^{-1/\alpha}.$$

(iv) If $\beta = 1$ and $(\alpha+1)m + d \neq 0$, then $\overline{G}$ satisfies condition $\mathcal{C}_2(\gamma, -1/\alpha, B_2)$, where

$$B_2(t) = \left(-\frac{(\alpha+1)m+d}{\alpha+1}(c/\alpha)^{-1/\alpha}\right) \times \frac{1}{\alpha} \times t^{-1/\alpha} = b_2\gamma t^{-1/\alpha}$$
$$\textit{with } b_2 = -\frac{(\alpha+1)m+d}{\alpha+1}(c/\alpha)^{-1/\alpha}.$$

(v) If $\beta = 1$ and $(\alpha+1)m + d = 0$, but $\beta' < 1$, then $\overline{G}$ satisfies condition $\mathcal{C}_2(\gamma, -(1+\beta')/\alpha, B_3)$, where

$$B_3(t) = \left(-\frac{d'(1+\beta')}{\alpha+\beta'+1}(c/\alpha)^{-(1+\beta')/\alpha}\right) \times \frac{1}{\alpha} \times t^{-(1+\beta')/\alpha} = b_3\gamma t^{-(1+\beta')/\alpha}$$
$$\textit{with } b_3 = -\frac{d'(1+\beta')}{\alpha+\beta'+1}(c/\alpha)^{-(1+\beta')/\alpha}.$$

(vi) If $\beta = 1$ and $(\alpha+1)m + d = 0$, but $\beta' > 1$, then $\overline{G}$ satisfies condition $\mathcal{C}_2(\gamma, -2/\alpha, B_4)$, where

$$B_4(t) = \frac{d^2}{\alpha+1}(c/\alpha)^{-2/\alpha} \times \frac{1}{\alpha} \times t^{-2/\alpha} = b_4\gamma t^{-2/\alpha}$$
$$\textit{with } b_4 = \frac{d^2}{\alpha+1}(c/\alpha)^{-2/\alpha}.$$

(vii) If $\beta = 1$, $(\alpha+1)m + d = 0$, $\beta' = 1$ and $2d'(\alpha+1) \neq d^2(\alpha+2)$, then $\overline{G}$ satisfies condition $\mathcal{C}_2(\gamma, -2/\alpha, B_5)$, where

$$B_5(t) = \left(-\frac{2}{\alpha+2}\left[d' - d^2\frac{\alpha+2}{2(\alpha+1)}\right](c/\alpha)^{-2/\alpha}\right) \times \frac{1}{\alpha} \times t^{-2/\alpha} = b_5\gamma t^{-2/\alpha}$$
$$\textit{with } b_5 = -\frac{2}{\alpha+2}\left[d' - d^2\frac{\alpha+2}{2(\alpha+1)}\right](c/\alpha)^{-2/\alpha}.$$

Heavy-tailed distributions with negative second-order parameter typically satisfy the third-order expansion

$$f(t) = ct^{-\alpha-1}(1 + dt^{-\beta} + d't^{-\beta-\beta'}(1 + \mathrm{o}(1))) \text{ as } t \to \infty$$

where $\alpha, \beta, \beta', c > 0$ and $d, d' \neq 0$, as well as (depending on m) one of the conditions in (ii)-(vii). The only exception among the usual families of heavy-tailed distributions seems to be the Generalized Pareto distribution, for which it can be the case that $f(t-m) = ct^{-\alpha-1}$ for suitable $\alpha, c > 0$, and thus no remainder term is present. This, however, represents the ideal case when the shifted distribution is exactly Pareto and no bias is incurred by the use of the heavy tail assumption.

Proof. The assumption

$$f(t) = ct^{-\alpha-1}(1 + dt^{-\beta}(1 + \mathrm{o}(1))) \text{ as } t \to \infty$$

yields

$$\overline{F}(t) = \frac{c}{\alpha}t^{-\alpha}\left(1 + \frac{d\alpha}{\alpha+\beta}t^{-\beta}(1 + \mathrm{o}(1))\right).$$

In particular, for any $x > 0$,

$$\frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-\alpha} = \left(-\frac{d\beta}{\alpha+\beta}\right)\frac{1}{\alpha}t^{-\beta} \times x^{-\alpha}\frac{x^{-\beta}-1}{(-\beta/\alpha)\times(1/\alpha)} + \mathrm{o}(t^{-\beta}).$$

It is then straightforward to check statement (i). The key to showing statements (ii)–(vii) is to note that when

$$f(t) = ct^{-\alpha-1}(1 + dt^{-\beta} + d't^{-\beta-\beta'}(1+\mathrm{o}(1))) \text{ as } t \to \infty,$$

it similarly holds that

$$\begin{aligned}
f(t-m) &= c(t-m)^{-\alpha-1}(1 + d(t-m)^{-\beta} + d't^{-\beta-\beta'}(1+\mathrm{o}(1))) \\
&= ct^{-\alpha-1}(1-mt^{-1})^{-\alpha-1}(1 + dt^{-\beta}(1-mt^{-1})^{-\beta} + d't^{-\beta-\beta'}(1+\mathrm{o}(1))) \\
&= ct^{-\alpha-1}\left(1 + (\alpha+1)mt^{-1} + \frac{(\alpha+1)(\alpha+2)m^2}{2}t^{-2} + \mathrm{o}(t^{-2})\right) \\
&\times (1 + dt^{-\beta} + \beta dmt^{-\beta-1} + d't^{-\beta-\beta'} + \mathrm{o}(t^{-\beta-1}) + \mathrm{o}(t^{-\beta-\beta'})) \\
&= ct^{-\alpha-1}\left(1 + (\alpha+1)mt^{-1} + dt^{-\beta} + \frac{(\alpha+1)(\alpha+2)m^2}{2}t^{-2} + (\alpha+\beta+1)dmt^{-\beta-1}\right. \\
&\left. + d't^{-\beta-\beta'} + \mathrm{o}(t^{-2}) + \mathrm{o}(t^{-\beta-1}) + \mathrm{o}(t^{-\beta-\beta'})\right).
\end{aligned}$$

We then find that, as $t \to \infty$:

- In the setup of (ii),
$$f(t-m) = ct^{-\alpha-1}(1 + dt^{-\beta}(1+\mathrm{o}(1))).$$
- In the setup of (iii),
$$f(t-m) = ct^{-\alpha-1}(1 + (\alpha+1)mt^{-1}(1+\mathrm{o}(1))).$$
- In the setup of (iv),
$$f(t-m) = ct^{-\alpha-1}(1 + ((\alpha+1)m + d)t^{-1}(1+\mathrm{o}(1))).$$
- In the setup of (v),
$$f(t-m) = ct^{-\alpha-1}(1 + d't^{-1-\beta'}(1+\mathrm{o}(1))).$$
- In the setup of (vi),
$$f(t-m) = ct^{-\alpha-1}\left(1 - \frac{d^2(\alpha+2)}{2(\alpha+1)}t^{-2}(1+\mathrm{o}(1))\right).$$
- In the setup of (vii),
$$f(t-m) = ct^{-\alpha-1}\left(1 + \left(d' - \frac{d^2(\alpha+2)}{2(\alpha+1)}\right)t^{-2}(1+\mathrm{o}(1))\right).$$

The result then immediately follows from (i). □

Distribution (parameters)	Density function	γ	ρ	b
Pareto ($\alpha > 0$)	$\alpha t^{-\alpha-1}$ $(t > 1)$	$1/\alpha$	$-\infty$	0 (by convention)
Hall-Weiss ($\alpha, \beta > 0$)	$(\alpha t^{-\alpha-1} + (\alpha+\beta)t^{-\alpha-\beta-1})/2$ $(t > 1)$	$1/\alpha$	$-\beta/\alpha$	$-\beta\, 2^{\beta/\alpha}/\alpha$
Burr ($\alpha, \beta, \lambda > 0$)	$\frac{\alpha\beta}{\lambda}\left(\frac{t}{\lambda}\right)^{\alpha-1}\left(1+\left(\frac{t}{\lambda}\right)^{\alpha}\right)^{-\beta-1}$ $(t > 0)$	$1/(\alpha\beta)$	$-1/\beta$	1
Dagum ($\alpha, \beta, \lambda > 0$)	$\frac{\alpha\beta}{\lambda}\left(\frac{t}{\lambda}\right)^{-\alpha-1}\left(1+\left(\frac{t}{\lambda}\right)^{-\alpha}\right)^{-\beta-1}$ $(t > 0)$	$1/\alpha$	-1	$\frac{\beta+1}{2\beta}$
Fréchet ($\alpha > 0$)	$\alpha t^{-\alpha-1}\exp\left(-t^{-\alpha}\right)$ $(t > 0)$	$1/\alpha$	-1	$1/2$
Generalized Pareto ($\sigma, \xi > 0$)	$\sigma^{-1}(1+\xi t/\sigma)^{-1-1/\xi}$ $(t > 0)$	ξ	$-\xi$	1
Fisher ($\nu_1, \nu_2 > 0$)	$\frac{(\nu_1/\nu_2)^{\nu_1/2}}{B(\nu_1/2, \nu_2/2)} t^{\nu_1/2-1}(1+\nu_1 t/\nu_2)^{-(\nu_1+\nu_2)/2}$ $(t > 0)$	$2/\nu_2$	$-2/\nu_2$	$\frac{\nu_1+\nu_2}{\nu_2+2}(B(\nu_1/2, \nu_2/2))^{2/\nu_2}(\nu_2/2)^{2/\nu_2}$
Lévy ($\lambda > 0$)	$\sqrt{\frac{\lambda}{2\pi}}t^{-3/2}\exp(-\lambda/(2t))$ $(t > 0)$	2	-2	$\frac{\pi}{6}$
Inverse-Γ ($\alpha, \lambda > 0$)	$\frac{\lambda^\alpha}{\Gamma(\alpha)}t^{-\alpha-1}\exp(-\lambda/t)$ $(t > 0)$	$1/\alpha$	$-1/\alpha$	$(\Gamma(\alpha+1))^{1/\alpha}/(\alpha+1)$
Davis ($\alpha > 1, \lambda > 0$)	$\frac{\lambda^\alpha t^{-\alpha-1}}{\Gamma(\alpha)\zeta(\alpha)(e^{\lambda/t}-1)}$ $(t > 0)$	$1/(\alpha-1)$	$-1/(\alpha-1)$	$\frac{1}{2\alpha}\{(\alpha-1)\Gamma(\alpha)\zeta(\alpha)\}^{1/(\alpha-1)}$
Slash ($\alpha > 0$)	$\frac{2^{\alpha/2-1}\alpha\Gamma\left(\frac{\alpha+1}{2}\right)}{\sqrt{\pi}}\chi^2_{1+\alpha}(t^2)\lvert t\rvert^{-1-\alpha}$ $(t \in \mathbb{R}\setminus\{0\})$	$1/\alpha$	$-\infty$	0 (by convention)
Cauchy ($c > 0$)	$\frac{c}{\pi(c^2+t^2)}$ $(t \in \mathbb{R})$	1	-2	$2\pi^2/3$
Student ($\nu > 0$)	$\frac{1}{\sqrt{\nu\pi}}\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$ $(t \in \mathbb{R})$	$1/\nu$	$-2/\nu$	$\frac{\nu(\nu+1)}{\nu+2}\left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)\nu^{(\nu-1)/2}}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\right)^{-2/\nu}$

Table C2: A list of standard continuous heavy-tailed distributions satisfying $\mathcal{C}_2(\gamma, \rho, A)$ with $A(t) = b\gamma t^\rho$, with the associated values of γ, ρ and b. The Burr distribution in this parametrization is also called the Singh-Maddala distribution, and it coincides with the log-logistic distribution when $\beta = 1$. The Dagum distribution is also called Burr type III distribution on p.59 of Beirlant et al. (2004). Γ stands for Euler's Gamma function, and in the density function of the Davis distribution, ζ denotes Riemann's Zeta function. In the Slash density function, $\chi^2_{1+\alpha}(\cdot)$ stands for the cumulative distribution function of the χ^2 distribution with $1+\alpha$ degrees of freedom.

The following lemma is a law of large numbers, with rate of convergence, for $\widehat{\overline{F}}_n$ and $\widehat{\varphi}_n^{(1)}$ at random thresholds. It is the crucial element for our proof of Theorem 1(i), with the quantification of the rate being a prerequisite for the proof of Lemma 3 below, which is itself key to the proof of Theorem 1(ii). Before that we recall the following fact: if condition $\mathcal{C}_2(\gamma, \rho, A)$ is satisfied and $\gamma < 1/k$, where k is a positive integer, then

$$\begin{aligned}\varphi^{(k)}(x) &= kx^k\overline{F}(x)\left(\int_1^\infty (v-1)^{k-1}v^{-1/\gamma}\,dv + \int_1^\infty (v-1)^{k-1}\left\{\frac{\overline{F}(vx)}{\overline{F}(x)} - v^{-1/\gamma}\right\}dv\right)\\ &= x^k\overline{F}(x)\left(\frac{1}{\gamma}B(k+1,\gamma^{-1}-k)\right.\\ &\left.+\frac{A(1/\overline{F}(x))}{\gamma\rho}\left\{\frac{1-\rho}{\gamma}B(k+1,\gamma^{-1}(1-\rho)-k) - \frac{1}{\gamma}B(k+1,\gamma^{-1}-k) + \mathrm{o}(1)\right\}\right)\end{aligned} \tag{C1}$$

as $x \to \infty$, using uniform second-order regular variation inequalities: see Theorem B.2.18 in de Haan and Ferreira (2006), which applies since second-order regular variation can be rewritten as extended regular variation (this is explicitly stated as Lemma 5 in Daouia et al. (2020)). In this equation, $B(\cdot,\cdot)$ denotes the Beta function. In particular,

$$\varphi^{(k)}(x) = \frac{B(k+1,\gamma^{-1}-k)}{\gamma}x^k\overline{F}(x)(1+\mathrm{O}(A(1/\overline{F}(x)))) \text{ as } x\to\infty. \tag{C2}$$

Lemma 2. *Assume that condition $\mathcal{C}_2(\gamma,\rho,A)$ is satisfied. Let $Y_1,\ldots,Y_n$ be independent random variables with distribution function F, and assume that $\tau_n \uparrow 1$ satisfies $n(1-\tau_n)\to\infty$ and $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) = \mathrm{O}(1)$ as $n\to\infty$. Let finally $\widehat{c}_n$ be a random sequence converging in probability to a fixed constant $c>0$ at the rate $1/\sqrt{n(1-\tau_n)}$, namely, $\widehat{c}_n - c = \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})$.*
(i) Then

$$\frac{\widehat{\overline{F}}_n(\widehat{c}_n q_{\tau_n})}{\overline{F}(cq_{\tau_n})} = \frac{1}{n}\sum_{i=1}^n \frac{\mathbb{1}\{Y_i > \widehat{c}_n q_{\tau_n}\}}{\mathbb{P}(Y > cq_{\tau_n})} = 1 + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}).$$

(ii) If $0<\gamma<1/2$ then

$$\frac{\widehat{\varphi}_n^{(1)}(\widehat{c}_n q_{\tau_n})}{\varphi^{(1)}(cq_{\tau_n})} = \frac{1}{n}\sum_{i=1}^n \frac{(Y_i - \widehat{c}_n q_{\tau_n})\mathbb{1}\{Y_i > \widehat{c}_n q_{\tau_n}\}}{\mathbb{E}((Y-cq_{\tau_n})\mathbb{1}\{Y > cq_{\tau_n}\})} = 1 + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}).$$

Proof of Lemma 2. (i) Obviously

$$\frac{\widehat{\overline{F}}_n(cq_{\tau_n})}{\overline{F}(cq_{\tau_n})} = 1 + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}). \tag{C3}$$

This is immediate because the left-hand side is a mean of independent and identically distributed random variables for each n, having mean 1, and

$$\mathrm{Var}\left(\frac{\widehat{\overline{F}}_n(cq_{\tau_n})}{\overline{F}(cq_{\tau_n})}\right) = \frac{1}{n}\times\frac{\mathrm{Var}(\mathbb{1}\{Y>cq_{\tau_n}\})}{(\overline{F}(cq_{\tau_n}))^2} = \frac{1}{n}\left(\frac{1}{\overline{F}(cq_{\tau_n})} - 1\right)$$

is asymptotically equivalent to a multiple of $1/(n(1-\tau_n))$, by the regular variation assumption on $\overline{F}$. It then suffices to check that

$$\frac{\widehat{\overline{F}}_n(\widehat{c}_n q_{\tau_n})}{\widehat{\overline{F}}_n(cq_{\tau_n})} = 1 + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}).$$

Write
$$\left|\frac{\widehat{\overline{F}}_n(\widehat{c}_n q_{\tau_n})}{\widehat{\overline{F}}_n(cq_{\tau_n})}-1\right| \le \frac{1+\mathrm{o}_{\mathbb{P}}(1)}{\overline{F}(cq_{\tau_n})}\times\frac{1}{n}\sum_{i=1}^{n}|\mathbb{1}\{Y_i>\widehat{c}_n q_{\tau_n}\}-\mathbb{1}\{Y_i>cq_{\tau_n}\}|.$$
It follows from $\widehat{c}_n - c = \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})$ and (C3) that, for any $\varepsilon>0$, there is $M>0$ such that, for n large enough, one has

$$\begin{aligned}
&\left|\frac{\widehat{\overline{F}}_n(\widehat{c}_n q_{\tau_n})}{\widehat{\overline{F}}_n(cq_{\tau_n})}-1\right| \\
&\le \frac{2}{\overline{F}(cq_{\tau_n})}\times\frac{1}{n}\sum_{i=1}^{n}(\mathbb{1}\{Y_i>c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n}\}-\mathbb{1}\{Y_i>c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n}\}) \\
&= \frac{2(\widehat{\overline{F}}_n(c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n})-\widehat{\overline{F}}_n(c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n}))}{\overline{F}(cq_{\tau_n})} \\
&= \frac{2(\overline{F}(c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n})-\overline{F}(c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n}))}{\overline{F}(cq_{\tau_n})}+\mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})
\end{aligned}$$

with probability larger than $1-\varepsilon/2$. Now

$$\frac{2(\overline{F}(c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n})-\overline{F}(c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n}))}{\overline{F}(cq_{\tau_n})}=\mathrm{O}(1/\sqrt{n(1-\tau_n)})$$

using the assumption $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1})=\mathrm{O}(1)$ and the locally uniform character of second-order regular variation (see again Theorem B.2.18 in de Haan and Ferreira (2006)). Conclude that there is $M'>0$ such that, for n large enough, one has

$$\sqrt{n(1-\tau_n)}\left|\frac{\widehat{\overline{F}}_n(\widehat{c}_n q_{\tau_n})}{\widehat{\overline{F}}_n(cq_{\tau_n})}-1\right|\le M'$$

with probability larger than $1-\varepsilon$. The result follows.

(ii) We mimic the proof of (i) with a couple of adaptations. First of all

$$\frac{\widehat{\varphi}_n^{(1)}(cq_{\tau_n})}{\varphi^{(1)}(cq_{\tau_n})}=1+\mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}). \tag{C4}$$

Indeed, the left-hand side is a mean of independent and identically distributed random variables for each n, having mean 1, and

$$\mathrm{Var}\left(\frac{\widehat{\varphi}_n^{(1)}(cq_{\tau_n})}{\varphi^{(1)}(cq_{\tau_n})}\right)=\frac{1}{n}\times\frac{\mathrm{Var}((Y-cq_{\tau_n})\mathbb{1}\{Y>cq_{\tau_n}\})}{(\varphi^{(1)}(cq_{\tau_n}))^2}=\frac{1}{n}\left(\frac{\varphi^{(2)}(cq_{\tau_n})}{(\varphi^{(1)}(cq_{\tau_n}))^2}-1\right)$$

which, by Lemma 3(i) in Stupfler and Usseglio-Carleve (2023), is asymptotically equivalent to a multiple of $1/(n(1-\tau_n))$. It is then enough to prove that

$$\frac{\widehat{\varphi}_n^{(1)}(\widehat{c}_n q_{\tau_n})}{\widehat{\varphi}_n^{(1)}(cq_{\tau_n})}=1+\mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}).$$

Note then that

$$
\begin{aligned}
&\frac{\widehat{\varphi}_n^{(1)}(\widehat{c}_n q_{\tau_n})}{\widehat{\varphi}_n^{(1)}(c q_{\tau_n})} - 1 \\
&= \frac{1+\mathrm{o}_{\mathbb{P}}(1)}{\varphi^{(1)}(c q_{\tau_n})} \times \frac{1}{n}\sum_{i=1}^n (Y_i - c q_{\tau_n})(\mathbb{1}\{Y_i > \widehat{c}_n q_{\tau_n}\} - \mathbb{1}\{Y_i > c q_{\tau_n}\}) \\
&- (\widehat{c}_n - c)\frac{q_{\tau_n}\overline{F}(c q_{\tau_n})}{\varphi^{(1)}(c q_{\tau_n})} \times \frac{\widehat{\overline{F}}_n(\widehat{c}_n q_{\tau_n})}{\overline{F}(c q_{\tau_n})}(1+\mathrm{o}_{\mathbb{P}}(1)) \\
&= \frac{1+\mathrm{o}_{\mathbb{P}}(1)}{\varphi^{(1)}(c q_{\tau_n})} \times \frac{1}{n}\sum_{i=1}^n (Y_i - c q_{\tau_n})(\mathbb{1}\{Y_i > \widehat{c}_n q_{\tau_n}\} - \mathbb{1}\{Y_i > c q_{\tau_n}\}) + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})
\end{aligned}
$$

by (i) of the present lemma and Lemma 3(i) in Stupfler and Usseglio-Carleve (2023). Then, for any $\varepsilon > 0$, there is $M > 0$ such that, for n large enough, one has

$$
\begin{aligned}
&\left|\frac{1}{\varphi^{(1)}(c q_{\tau_n})} \times \frac{1}{n}\sum_{i=1}^n (Y_i - c q_{\tau_n})(\mathbb{1}\{Y_i > \widehat{c}_n q_{\tau_n}\} - \mathbb{1}\{Y_i > c q_{\tau_n}\})\right| \\
&\le \frac{1}{\varphi^{(1)}(c q_{\tau_n})} \times \frac{1}{n}\sum_{i=1}^n |Y_i - c q_{\tau_n}|(\mathbb{1}\{Y_i > c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n}\} - \mathbb{1}\{Y_i > c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n}\}) \\
&\le \frac{1}{\varphi^{(1)}(c q_{\tau_n})}\left(\frac{1}{n}\sum_{i=1}^n (Y_i - c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n})\mathbb{1}\{Y_i > c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n}\}\right. \\
&\qquad - \frac{1}{n}\sum_{i=1}^n (Y_i - c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n})\mathbb{1}\{Y_i > c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n}\} \\
&\left. + \frac{M}{\sqrt{n(1-\tau_n)}} c q_{\tau_n} \times \frac{1}{n}\sum_{i=1}^n (\mathbb{1}\{Y_i > c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n}\} - \mathbb{1}\{Y_i > c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n}\})\right) \\
&= \frac{\varphi^{(1)}(c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n})}{\varphi^{(1)}(c q_{\tau_n})} - \frac{\varphi^{(1)}(c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n})}{\varphi^{(1)}(c q_{\tau_n})} \\
&+ \frac{M}{\sqrt{n(1-\tau_n)}} c q_{\tau_n}\left\{\frac{\overline{F}(c(1-M/\sqrt{n(1-\tau_n)})q_{\tau_n})}{\varphi^{(1)}(c q_{\tau_n})} - \frac{\overline{F}(c(1+M/\sqrt{n(1-\tau_n)})q_{\tau_n})}{\varphi^{(1)}(c q_{\tau_n})}\right\} \\
&+ \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})
\end{aligned}
$$

with probability larger than $1 - \varepsilon/2$ (where (C3) and (C4) were used again). Using (C2) and again the locally uniform character of second-order regular variation, we find that the upper bound is a $\mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})$, thus completing the proof. □

Our final lemma quantifies the error made in the estimation of the various matrices introduced in the construction of our refined confidence interval for the LAWS estimator.

Lemma 3. *Assume that $\mathbb{E}(|\min(Y,0)|^{2+\delta}) < \infty$ for some $\delta > 0$, and condition $\mathcal{C}_2(\gamma,\rho,A)$ is satisfied with $0 < \gamma < 1/2$, $\rho < 0$ and $A(t) = b\gamma t^\rho$. Let $\tau_n \uparrow 1$ satisfy $n(1-\tau_n) \to \infty$ and $\sqrt{n(1-\tau_n)}(1/q_{\tau_n} + A((1-\tau_n)^{-1})) = \mathrm{O}(1)$ as $n \to \infty$. Throughout this lemma, let $\widehat{\boldsymbol{M}}_n^\varphi$, $\widehat{\boldsymbol{M}}_n^E$, $\widehat{\boldsymbol{M}}_n^\xi$, $\widehat{\boldsymbol{M}}_n$ and $\widehat{\mathfrak{M}}_n(J)$ be the random matrices deduced from $\boldsymbol{M}_n^\varphi$, $\boldsymbol{M}_n^E$, $\boldsymbol{M}_n^\xi$, $\boldsymbol{M}_n$ and $\mathfrak{M}_n(J)$, respectively, by replacing γ by $\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}$, ξ_{τ_n} by $\widehat{\xi}_{\tau_n}$, $\mathbb{E}(Y)$ by $\overline{Y}_n$, $\overline{F}(\xi_{\tau_n})$ by $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$, $\varphi^{(1)}(\xi_{\tau_n})$ by $\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n})$ and $\varphi^{(2)}(\xi_{\tau_n})$ by $\widetilde{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n})$ (where in the latter, $\overline{\rho}$ and $\overline{b}$ are consistent estimators of ρ and b such that $(\overline{\rho}-\rho)\log(n) = \mathrm{o}_{\mathbb{P}}(1)$), where throughout*

$\alpha_n = F(\xi_{\tau_n})$. *Then, elementwise, for any* $J \geq 1$,

$$\widehat{\boldsymbol{M}}_n^E = \widehat{\boldsymbol{M}}_n^\varphi + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})$$
$$= \begin{pmatrix} \dfrac{2\gamma}{1-2\gamma} & \dfrac{\gamma}{1-\gamma} \\ \dfrac{\gamma}{1-\gamma} & \dfrac{\gamma}{1-\gamma} \end{pmatrix} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}), \tag{C5}$$

$$\widehat{\boldsymbol{M}}_n^\xi = \begin{pmatrix} \dfrac{2\gamma^3}{1-2\gamma} & \dfrac{\gamma^3}{1-\gamma} \\ \dfrac{\gamma^3}{1-\gamma} & \dfrac{\gamma^3}{1-\gamma} \end{pmatrix} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}), \tag{C6}$$

$$\widehat{\boldsymbol{M}}_n = \begin{pmatrix} \dfrac{1-\gamma}{\gamma(1-2\gamma)} & -\dfrac{\gamma}{1-2\gamma} \\ -\dfrac{\gamma}{1-2\gamma} & \dfrac{2\gamma^3}{1-2\gamma} \end{pmatrix} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}) \tag{C7}$$

$$\textit{and } \widehat{\mathfrak{M}}_n(J) = \begin{pmatrix} \dfrac{\gamma^3(1-\gamma)}{(1-2\gamma)} & \dfrac{\gamma^3}{1-2\gamma} \\ \dfrac{\gamma^3}{1-2\gamma} & \dfrac{2\gamma^3}{1-2\gamma} \end{pmatrix} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}). \tag{C8}$$

Proof of Lemma 3. Under the stated assumptions:

- $\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}$ is a $\sqrt{n(1-\tau_n)}-$consistent estimator of γ, see Theorem 1 in Girard et al. (2022).
- $\widehat{\xi}_{\tau_n}$ is a $\sqrt{n(1-\tau_n)}-$relatively consistent estimator of ξ_{τ_n} in the sense that $\widehat{\xi}_{\tau_n}/\xi_{\tau_n} - 1 = \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})$, see Theorem 2 in Daouia et al. (2018) or Theorem 1 in Daouia et al. (2020).
- $\overline{Y}_n$ is a $\sqrt{n}-$consistent estimator of $\mathbb{E}(Y)$ by the standard central limit theorem.
- $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$ and $\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n})$ are $\sqrt{n(1-\tau_n)}-$relatively consistent estimators of $\overline{F}(\xi_{\tau_n})$ and $\varphi^{(1)}(\xi_{\tau_n})$, see Lemma 2.
- $\widetilde{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n})$ is a $\sqrt{n(1-\tau_n)}-$relatively consistent estimator of $\varphi^{(2)}(\xi_{\tau_n})$, because

$$\widetilde{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n}) = \frac{2\gamma^2}{(1-\gamma)(1-2\gamma)}\xi_{\tau_n}^2\overline{F}(\xi_{\tau_n})(1+\mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}))$$
$$= \varphi^{(2)}(\xi_{\tau_n})(1+\mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})),$$

see Equation (C2).

These convergences will be used without further mention throughout the proof. We start by noting that

$$\frac{\widehat{\xi}_{\tau_n} - \overline{Y}_n}{2\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n}) + \widehat{\xi}_{\tau_n} - \overline{Y}_n} - 1 = -\frac{2\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n})}{2\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n}) + \widehat{\xi}_{\tau_n} - \overline{Y}_n} = -2\frac{\varphi^{(1)}(\xi_{\tau_n})}{\xi_{\tau_n}}(1+\mathrm{o}_{\mathbb{P}}(1))$$

by Lemma 2. By Equation (C2) then,

$$\frac{\widehat{\xi}_{\tau_n} - \overline{Y}_n}{2\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n}) + \widehat{\xi}_{\tau_n} - \overline{Y}_n} - 1 = \mathrm{O}_{\mathbb{P}}(\overline{F}(\xi_{\tau_n})) = \mathrm{O}_{\mathbb{P}}(1-\tau_n) = \mathrm{o}_{\mathbb{P}}(1/q_{\tau_n}) = \mathrm{o}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})$$

because q_{τ_n} is asymptotically equivalent to a multiple of $(1-\tau_n)^{-\gamma}$ (see the remark below Example 2.3.11 p.49 in de Haan and Ferreira (2006)), and $\gamma < 1$. Consequently $\widehat{M}^E_{n,11} = \widehat{M}^\varphi_{n,11}(1 + \mathrm{o}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}))$ and likewise $\widehat{M}^E_{n,12} = \widehat{M}^\varphi_{n,12}(1+\mathrm{o}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}))$. Now, since $1-\alpha_n = \overline{F}(\xi_{\tau_n})$ is estimated by $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$,

$$\widehat{M}^\varphi_{n,11} = (1-\tau_n)\left(\frac{\widetilde{\varphi}^{(2)}_n(\widehat{\xi}_{\tau_n})}{[\widehat{\varphi}^{(1)}_n(\widehat{\xi}_{\tau_n})]^2} - 1\right) \text{ and } \widehat{M}^\varphi_{n,12} = \widehat{M}^\varphi_{n,22} = (1-\tau_n)\frac{1-\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}.$$

Using Equation (C2) and assumption $\sqrt{n(1-\tau_n)}(1/q_{\tau_n} + A((1-\tau_n)^{-1})) = \mathrm{O}(1)$, this entails

$$\widehat{M}^\varphi_{n,11} = \frac{2\gamma}{1-2\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}) \text{ and } \widehat{M}^\varphi_{n,12} = \widehat{M}^\varphi_{n,22} = \frac{\gamma}{1-\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}).$$

This proves Equation (C5). Then, similarly,

$$\frac{\widehat{\varphi}^{(1)}_n(\widehat{\xi}_{\tau_n})(2\widehat{\varphi}^{(1)}_n(\widehat{\xi}_{\tau_n}) + \widehat{\xi}_{\tau_n} - \overline{Y}_n)}{\widehat{\xi}_{\tau_n}(\widehat{\varphi}^{(1)}_n(\widehat{\xi}_{\tau_n}) + \widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})(\widehat{\xi}_{\tau_n} - \overline{Y}_n))} = \gamma + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})$$

because of Equation (C2), and therefore

$$\widehat{M}^\xi_{n,11} = \frac{2\gamma^3}{1-2\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}),\ \widehat{M}^\xi_{n,12} = \frac{\gamma^3}{1-\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})$$
$$\text{and } \widehat{M}^\xi_{n,22} = \frac{\gamma^3}{1-\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}).$$

This proves Equation (C6). As a consequence

$$\begin{aligned}
\widehat{M}_{n,11} &= \frac{1}{(\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})^2}\left(\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}\right)^2 \left[\widehat{M}^\xi_{n,11} - 2\widehat{M}^\xi_{n,12} + \widehat{M}^\xi_{n,22}\right] \\
&= \frac{1-\gamma}{\gamma(1-2\gamma)} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}), \\
\widehat{M}_{n,12} &= \frac{1}{\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}}\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}\left[\widehat{M}^\xi_{n,12} - \widehat{M}^\xi_{n,11}\right] = -\frac{\gamma}{1-2\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}) \\
\text{and } \widehat{M}_{n,22} &= \widehat{M}^\xi_{n,11} = \frac{2\gamma^3}{1-2\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)})
\end{aligned}$$

which is Equation (C7). Finally, $\widehat{\mathfrak{M}}_{n,22}(J) = \widehat{M}_{n,22}$, and besides

$$\begin{aligned}
\widehat{\mathfrak{M}}_{n,11}(J) &= \left(1 + \frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}\right)^{-4}\widehat{M}_{n,11}(1 + \mathrm{O}_{\mathbb{P}}(1/(n(1-\tau_n)))) \\
&= \frac{\gamma^3(1-\gamma)}{1-2\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}) \\
\text{and } \widehat{\mathfrak{M}}_{n,12}(J) &= -\left(1 + \frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}\right)^{-2}\widehat{M}_{n,12}(1 + \mathrm{O}_{\mathbb{P}}(1/(n(1-\tau_n))))
\end{aligned}$$

$$= \frac{\gamma^3}{1-2\gamma} + \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}).$$

This proves Equation (C8) and completes the proof. □

C.2 Proofs of the main results

Proof of Theorem 1. Recall from the proof of Lemma 3 that one has, under the stated assumptions:

- $\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}$ is a $\sqrt{n(1-\tau_n)}$−consistent estimator of γ.
- $\widehat{\xi}_{\tau_n}$ is a $\sqrt{n(1-\tau_n)}$−relatively consistent estimator of ξ_{τ_n}.
- $\overline{Y}_n$ is a $\sqrt{n}$−consistent estimator of $\mathbb{E}(Y)$.
- $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$ and $\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n})$ are $\sqrt{n(1-\tau_n)}$−relatively consistent estimators of $\overline{F}(\xi_{\tau_n})$ and $\varphi^{(1)}(\xi_{\tau_n})$.
- $\widetilde{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n})$ is a $\sqrt{n(1-\tau_n)}$−relatively consistent estimator of $\varphi^{(2)}(\xi_{\tau_n})$.

We turn to the proof of the desired results.

(i) First of all

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))} \log\left(\frac{\widehat{\xi}_{\tau_n'}^{\star,\mathrm{BR}}}{\xi_{\tau_n'}}\right) \stackrel{d}{\longrightarrow} \mathcal{N}\left(0, \frac{\gamma^3(1-\gamma)}{1-2\gamma}\right)$$

by Theorem 2(i) in Girard et al. (2022). By Slutsky's lemma, it follows that it is sufficient to prove that $\widehat{s}_n^2(J) \to \gamma^3(1-\gamma)/(1-2\gamma)$ in probability. Now, from straightforward calculations,

$$\partial_1 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}, 1) \to 1 \text{ and } \partial_2 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}, 1) \to 0$$

in probability as $n \to \infty$, where $\widehat{u}_n$ is obtained from the function u_n by replacing γ by $\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}$, ξ_{τ_n} by $\widehat{\xi}_{\tau_n}$, $\mathbb{E}(Y)$ by $\overline{Y}_n$, $\overline{F}(\xi_{\tau_n})$ by $\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})$, $\varphi^{(1)}(\xi_{\tau_n})$ by $\widehat{\varphi}_n^{(1)}(\widehat{\xi}_{\tau_n})$ and $\varphi^{(2)}(\xi_{\tau_n})$ by $\widetilde{\varphi}_n^{(2)}(\widehat{\xi}_{\tau_n})$. Consequently, by Lemma 3,

$$\begin{aligned}
\widehat{\mathfrak{M}}_{n,11}^{\mathrm{BR}}(J) &= (\partial_1 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}, 1))^2 \widehat{\mathfrak{M}}_{n,11}(J) \\
&+ (\partial_2 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}, 1))^2 \widehat{\mathfrak{M}}_{n,22}(J) \\
&+ 2\partial_1 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}, 1) \partial_2 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}, 1) \widehat{\mathfrak{M}}_{n,12}(J), \\
&\to \frac{\gamma^3(1-\gamma)}{1-2\gamma}, \\
\widehat{\mathfrak{M}}_{n,12}^{\mathrm{BR}}(J) &= \partial_1 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}, 1) \widehat{\mathfrak{M}}_{n,12}(J) \\
&+ \partial_2 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1}, 1) \widehat{\mathfrak{M}}_{n,22}(J), \\
&\to \frac{\gamma^3}{1-2\gamma}, \text{ and} \\
\widehat{\mathfrak{M}}_{n,22}^{\mathrm{BR}}(J) &\to \frac{2\gamma^3}{1-2\gamma}
\end{aligned}$$

in probability as $n \to \infty$, with the same notational convention. With that convention again,

$$\partial_1 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1) \to 1 \text{ and } \partial_2 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1) \to 0$$

in probability as $n \to \infty$, so that

$$\widehat{s}_n^2(J) = (\partial_1 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1))^2 \widehat{\mathfrak{M}}_{n,11}^{\mathrm{BR}}(J) + 2\partial_1 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1) \partial_2 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1) \widehat{\mathfrak{M}}_{n,12}^{\mathrm{BR}}(J)$$

$$+ (\partial_2 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1))^2 \widehat{\mathfrak{M}}_{n,22}^{\mathrm{BR}}(J) \to \frac{\gamma^3(1-\gamma)}{1-2\gamma}$$

in probability as $n \to \infty$, as required.

(ii) Observe that

$$\exp\left(\pm \frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}} \sqrt{\widehat{s}_n^2(J)} \times z_{1-\alpha/2} \right) - \exp\left(\pm \frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}} \sqrt{\frac{(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})^3(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})}{1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}}} \times z_{1-\alpha/2} \right)$$

is, in probability, asymptotically equivalent to

$$\pm \frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}} \left(\sqrt{\widehat{s}_n^2(J)} - \sqrt{\frac{(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})^3(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})}{1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}}} \right) \times z_{1-\alpha/2},$$

and therefore also to

$$\pm \frac{1}{2} \sqrt{\frac{1-2\gamma}{\gamma^3(1-\gamma)}} \frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}} \left(\widehat{s}_n^2(J) - \frac{(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})^3(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})}{1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}} \right) \times z_{1-\alpha/2}.$$

This means that

$$\begin{aligned} &\mathrm{length}(\widehat{I}_{\tau_n'}^{(2)}(\alpha; J)) - \mathrm{length}(\widehat{I}_{\tau_n'}^{(1)}(\alpha)) \\ &= \widehat{\xi}_{\tau_n'}^{\star,\mathrm{BR}} \sqrt{\frac{1-2\gamma}{\gamma^3(1-\gamma)}} \frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}} \left(\widehat{s}_n^2(J) - \frac{(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})^3(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})}{1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}} \right) z_{1-\alpha/2}(1+\mathrm{o}_{\mathbb{P}}(1)). \end{aligned}$$

Now

$$\begin{aligned} \widehat{s}_n^2(J) - \frac{(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})^3(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})}{1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}} &= ((\partial_1 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1))^2 - 1) \widehat{\mathfrak{M}}_{n,11}^{\mathrm{BR}}(J) \\ &+ \widehat{\mathfrak{M}}_{n,11}^{\mathrm{BR}}(J) - \frac{(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})^3(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})}{1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}} + 2\partial_1 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1) \partial_2 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1) \widehat{\mathfrak{M}}_{n,12}^{\mathrm{BR}}(J) \\ &+ (\partial_2 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1))^2 \widehat{\mathfrak{M}}_{n,22}^{\mathrm{BR}}(J). \end{aligned}$$

Equation (B10) shows that $1 - \partial_1 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1) = \mathrm{o}_{\mathbb{P}}(1/\log((1-\tau_n)/(1-\tau_n')))$, and Equation (B11) shows that $\partial_2 \widehat{g}_n(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}, 1)$ is asymptotically equivalent to $1/\log((1-\tau_n)/(1-\tau_n'))$ in probability. Therefore

$$\begin{aligned} \widehat{s}_n^2(J) - \frac{(\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})^3(1-\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}})}{1-2\widehat{\gamma}_{\tau_n}^{\mathrm{E,BR}}} &= \widehat{\mathfrak{M}}_{n,11}^{\mathrm{BR}}(J) - \frac{\gamma^3(1-\gamma)}{1-2\gamma} \\ &+ \frac{1}{\log((1-\tau_n)/(1-\tau_n'))} \left(\frac{2\gamma^3}{1-2\gamma} + \mathrm{o}_{\mathbb{P}}(1) \right). \end{aligned} \quad \text{(C9)}$$

Besides

$$\begin{aligned}
&\widehat{\mathfrak{M}}^{\mathrm{BR}}_{n,11}(J) - \frac{\gamma^3(1-\gamma)}{1-2\gamma} \\
&= ((\partial_1 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1},1))^2 - 1)\widehat{\mathfrak{M}}_{n,11}(J) + \widehat{\mathfrak{M}}_{n,11}(J) - \frac{\gamma^3(1-\gamma)}{1-2\gamma} \\
&+ (\partial_2 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1},1))^2 \widehat{\mathfrak{M}}_{n,22}(J) \\
&+ 2\partial_1 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1},1)\partial_2 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1},1)\widehat{\mathfrak{M}}_{n,12}(J).
\end{aligned}$$

By Equations (B8) and (B9), $1-\partial_1 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1},1) = \mathrm{O}_{\mathbb{P}}(1/\xi_{\tau_n}) + \mathrm{O}_{\mathbb{P}}(1-\tau_n) = \mathrm{O}_{\mathbb{P}}(1/q_{\tau_n}) = \mathrm{O}_{\mathbb{P}}(1/\sqrt{n(1-\tau_n)}) = \mathrm{o}_{\mathbb{P}}(1/\log((1-\tau_n)/(1-\tau_n')))$, and similarly $\partial_2 \widehat{u}_n((1+\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})/(1-\tau_n))^{-1},1) = \mathrm{O}_{\mathbb{P}}(1/\xi_{\tau_n}) = \mathrm{o}_{\mathbb{P}}(1/\log((1-\tau_n)/(1-\tau_n')))$. Applying Lemma 3 thus yields

$$\widehat{\mathfrak{M}}^{\mathrm{BR}}_{n,11}(J) - \frac{\gamma^3(1-\gamma)}{1-2\gamma} = \mathrm{o}_{\mathbb{P}}(1/\log((1-\tau_n)/(1-\tau_n'))). \tag{C10}$$

Combining Equations (C9) and (C10) results in

$$\widehat{s}^2_n(J) - \frac{(\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})^3(1-\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n})}{1-2\widehat{\gamma}^{\mathrm{E,BR}}_{\tau_n}} = \frac{1}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{2\gamma^3}{1-2\gamma} + \mathrm{o}_{\mathbb{P}}(1)\right).$$

From this one can conclude that

$$\begin{aligned}
\mathrm{length}(\widehat{I}^{(2)}_{\tau_n'}(\alpha;J)) - \mathrm{length}(\widehat{I}^{(1)}_{\tau_n'}(\alpha)) &= 2 \times \frac{\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\sqrt{n(1-\tau_n)}}\left(\sqrt{\frac{\gamma^3}{(1-\gamma)(1-2\gamma)}} z_{1-\alpha/2} + \mathrm{o}_{\mathbb{P}}(1)\right) \\
&= 2 \times \frac{\xi_{\tau_n'}}{\sqrt{n(1-\tau_n)}}\left(\sqrt{\frac{\gamma^3}{(1-\gamma)(1-2\gamma)}} z_{1-\alpha/2} + \mathrm{o}_{\mathbb{P}}(1)\right).
\end{aligned}$$

This completes the proof. □

Proof of Theorem 2. The proof generally follows the ideas of the proof of Theorem 1 with the necessary adaptations. We discuss the proof in detail below.

(i) First of all

$$\frac{\sqrt{k_n}}{\log(k_n/(n(1-\tau_n')))} \log\left(\frac{\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\xi_{\tau_n'}}\right) \xrightarrow{d} \mathcal{N}(0,\gamma^2)$$

by Theorem 2(ii) in Girard et al. (2022). By Slutsky's lemma, it follows that it is sufficient to prove that $\widetilde{\sigma}^2_n(J) \to \gamma^2$ in probability. To this end note that each term $v_{j,12}$ and $v_{j,22}$ appearing in the expressions of $V_{12}(J)$ and $V_{22}(J)$ is a continuous function of $\gamma \in (0,1)$, because $\phi : x \mapsto -x\log(x^{-1}-1)$ is infinitely differentiable on $(0,1)$. As a consequence, the estimators $\widehat{V}_{12}(J)$ and $\widehat{V}_{22}(J)$ appearing in $\widetilde{\sigma}^2_n(J)$ converge in probability to $V_{12}(0) = V_{12} = \gamma^2 m(\gamma)$ and $V_{22}(0) = V_{22} = \gamma^2(1+(m(\gamma))^2)$, respectively. Moreover, $\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n}$ is a $\sqrt{k_n}-$consistent estimator of γ under the stated assumptions (as recalled in *e.g.* Girard et al. (2022)), $\overline{Y}_n$ is a consistent estimator of $\mathbb{E}(Y)$ by the weak law of large numbers, and

$$\log\left[\frac{1}{\xi_{\tau_n'}}\left(\frac{n(1-\tau_n')}{k_n}\right)^{-\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n}} (1/\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n} - 1)^{-\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n}} \widehat{q}_{1-k_n/n}\right] \to 0$$

in probability, using the asymptotic proportionality relationship $\xi_{\tau_n'} \sim (\gamma^{-1}-1)^{-\gamma} q_{\tau_n'}$ and the consistency of the Weissman extreme quantile estimator. The estimators $\partial_1 \widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)$ and $\partial_2 \widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)$ in $\widetilde{\sigma}_n^2(J)$ thus converge in probability to 1 and 0, respectively, by (B16) and (B17). Conclude that $\widetilde{\sigma}_n^2(J) \to \gamma^2$ in probability, as required.

(ii) The quantity

$$\exp\left(\pm \frac{\log(k_n/(n(1-\tau_n')))}{\sqrt{k_n}} \sqrt{\widetilde{\sigma}_n^2(J)} \times z_{1-\alpha/2}\right) - \exp\left(\pm \frac{\log(k_n/(n(1-\tau_n')))}{\sqrt{k_n}} \widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n} \times z_{1-\alpha/2}\right)$$

is, in probability, asymptotically equivalent to

$$\pm \frac{1}{2\gamma} \frac{\log(k_n/(n(1-\tau_n')))}{\sqrt{k_n}} (\widetilde{\sigma}_n^2(J) - (\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n})^2) \times z_{1-\alpha/2}.$$

Then clearly

$$\begin{aligned}
&\widetilde{\sigma}_n^2(J) - (\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n})^2 \\
&= (\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n})^2((\partial_1 \widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1))^2 - 1) + (\partial_2 \widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1))^2 \widehat{V}_{22}(J) \\
&+ 2\partial_1 \widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)\partial_2 \widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)\widehat{V}_{12}(J)
\end{aligned}$$

and it suffices to control this expression to obtain the asymptotic behavior of the difference in length of the two asymptotic confidence intervals. As in the proof of Theorem 1(ii), Equations (B16) and (B17) show that $1-\partial_1 \widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1) = \mathrm{o}_{\mathbb{P}}(1/\log(k_n/(n(1-\tau_n'))))$ and $\partial_2 \widehat{h}_n(\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n},1)$ is asymptotically equivalent (in probability) to $1/\log(k_n/(n(1-\tau_n')))$. As a consequence

$$\widetilde{\sigma}_n^2(J) - (\widehat{\gamma}^{\mathrm{H,BR}}_{1-k_n/n})^2 = 2\gamma^2 m(\gamma)/\log(k_n/(n(1-\tau_n'))) + \mathrm{o}_{\mathbb{P}}(1/\log(k_n/(n(1-\tau_n')))).$$

Conclude that

$$\begin{aligned}
\mathrm{length}(\widetilde{I}^{(2)}_{\tau_n'}(\alpha;J)) - \mathrm{length}(\widetilde{I}^{(1)}_{\tau_n'}(\alpha)) &= 2 \times \frac{\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}}{\sqrt{k_n}}(\gamma m(\gamma) z_{1-\alpha/2} + \mathrm{o}_{\mathbb{P}}(1)) \\
&= 2 \times \frac{\xi_{\tau_n'}}{\sqrt{k_n}}(\gamma m(\gamma) z_{1-\alpha/2} + \mathrm{o}_{\mathbb{P}}(1)).
\end{aligned}$$

The proof is complete. □

Appendix D Further finite-sample illustrations and results

D.1 Simulated data

This section contains a complete set of results in our 18 models, for both of our inference methods (LAWS and quantile-based) and our three sample sizes $n = 200, 1{,}000, 5{,}000$.

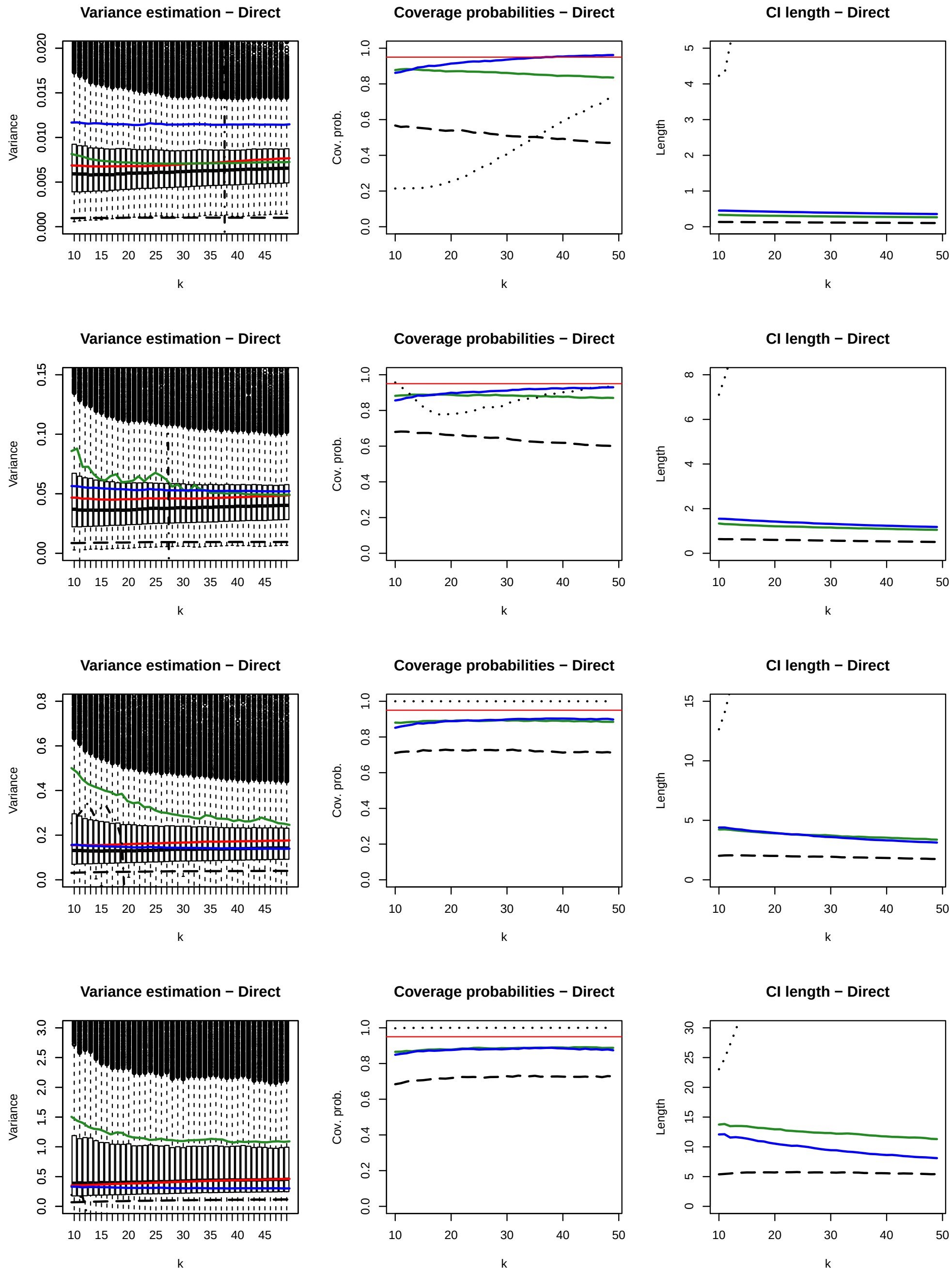

Fig. D1: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widehat{\xi}^{\star,\text{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widehat{s}_n^2(1)$ and the average values of $\widehat{s}_n^{2,\text{BR}}$ (dashed curve), $\widehat{s}_n^2(1)$ (green curve), $\widehat{s}_n^2$ (dashed-dotted curve) and $\overline{s}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha;1)$, $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widehat{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Burr distribution with $\rho = -5$ and $\gamma = 0.1$ (first row), $\gamma = 0.2$ (second row), $\gamma = 0.3$ (third row), $\gamma = 0.4$ (fourth row). The sample size is $n = 200$ and the target expectile level is $\tau_n' = 1 - 1/n = 0.995$.

D.2 Real data

This section contains extra results about our cyber insurance real data analysis. We provide estimates of the tail index (Figure D29), extreme quantile (Figure D30) and extreme expectile (Figure D31) in each cluster formed by stratifying upon the type of cyber breach.

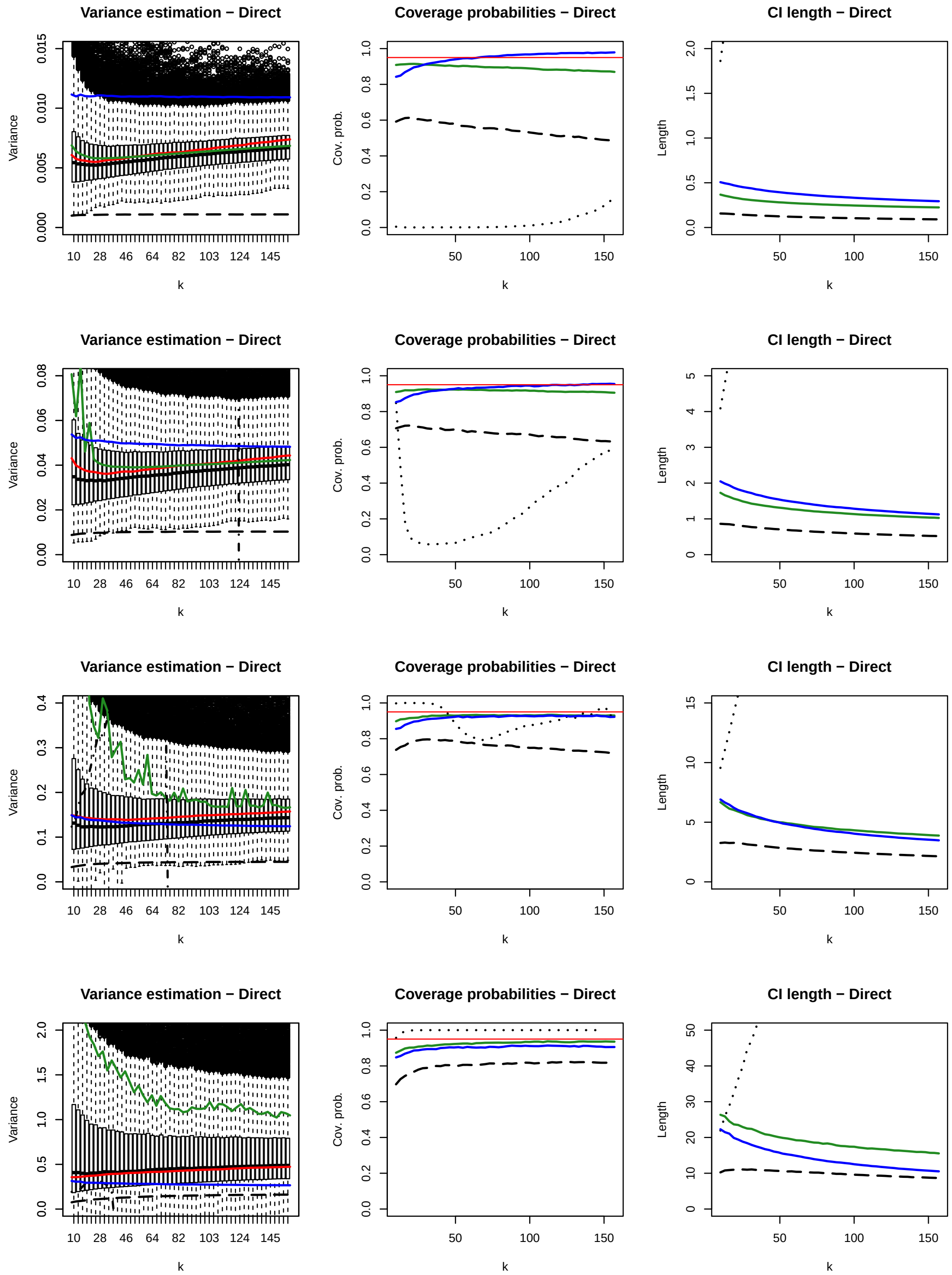

Fig. D2: As in Figure D1, with $n = 1{,}000$ and $\tau_n' = 1 - 1/n = 0.999$.

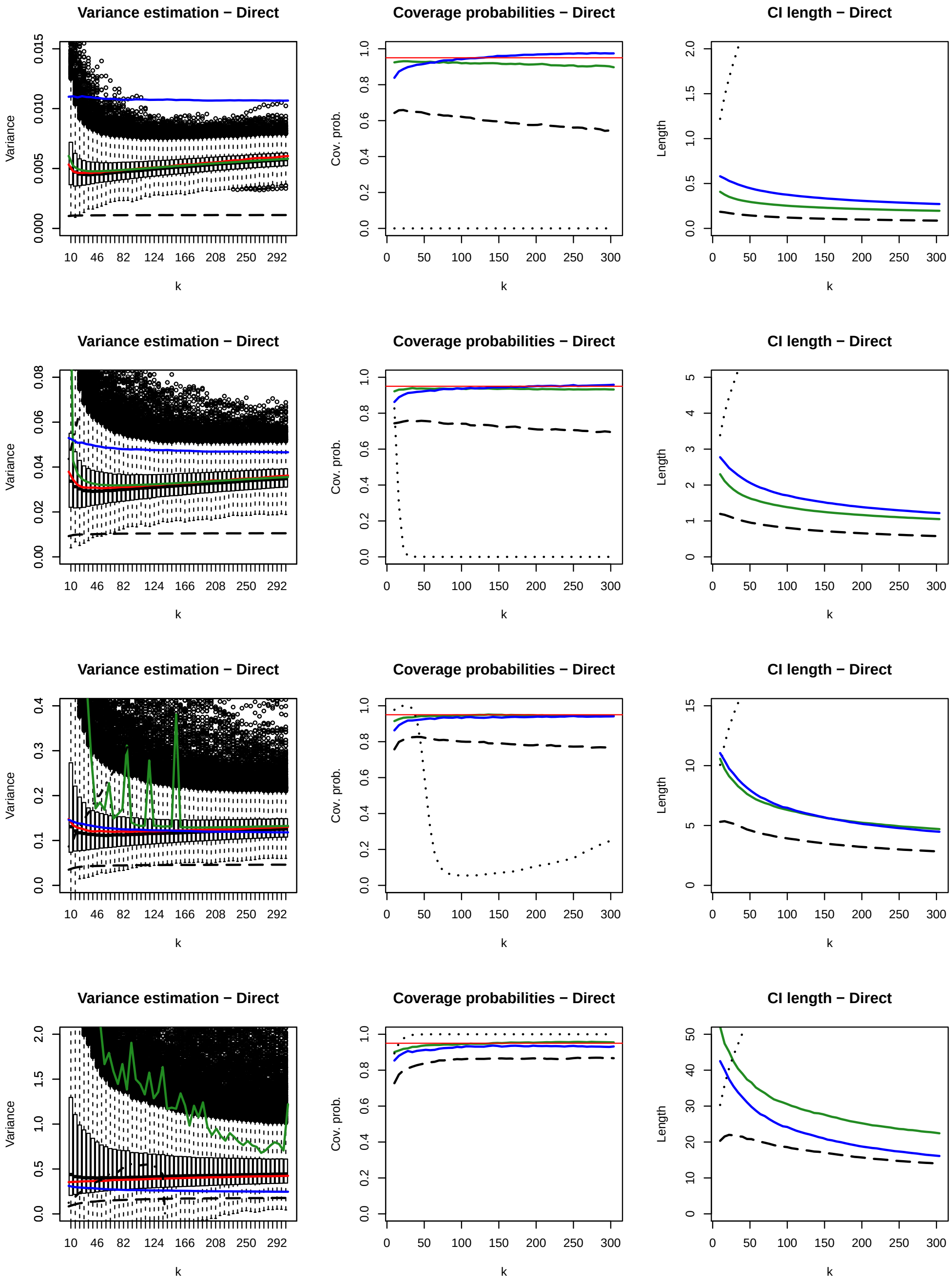

Fig. D3: As in Figure D1, with $n = 5{,}000$ and $\tau_n' = 1 - 1/n = 0.9998$.

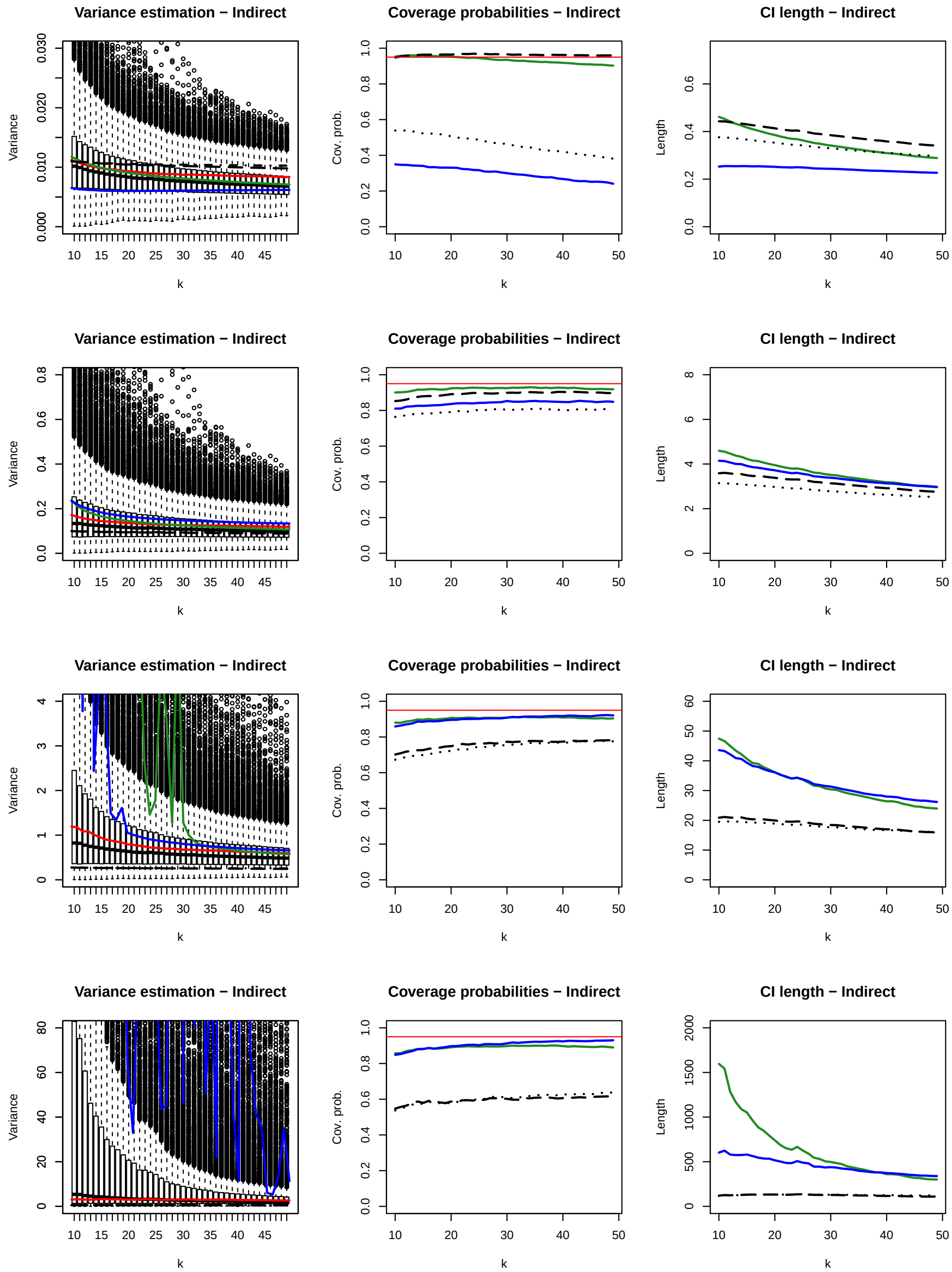

Fig. D4: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widetilde{\sigma}_n^2(2)$ and the average values of $\widetilde{\sigma}_n^{2,\mathrm{BR}}$ (dashed curve), $\widetilde{\sigma}_n^2(2)$ (green curve), $\widetilde{\sigma}_n^2$ (dashed-dotted curve) and $\overline{\sigma}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(2)}_{\tau_n'}(\alpha;2)$, $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Burr distribution with $\rho = -5$ and $\gamma = 0.1$ (first row), $\gamma = 0.3$ (second row), $\gamma = 0.5$ (third row), $\gamma = 0.7$ (fourth row). The sample size is $n = 200$ and the target expectile level is $\tau_n' = 1 - 1/n = 0.995$.

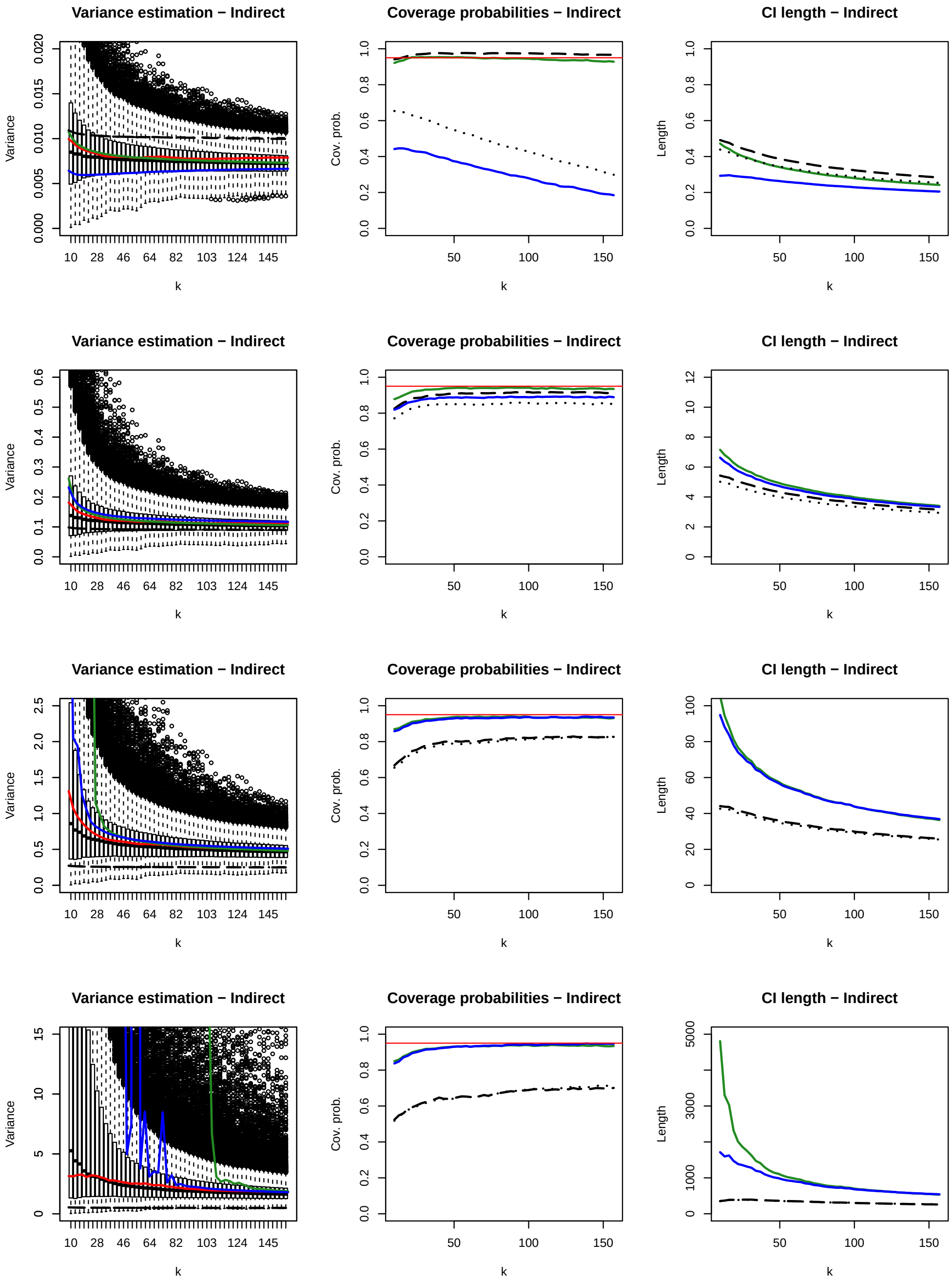

Fig. D5: As in Figure D4, with $n = 1{,}000$ and $\tau_n' = 1 - 1/n = 0.999$.

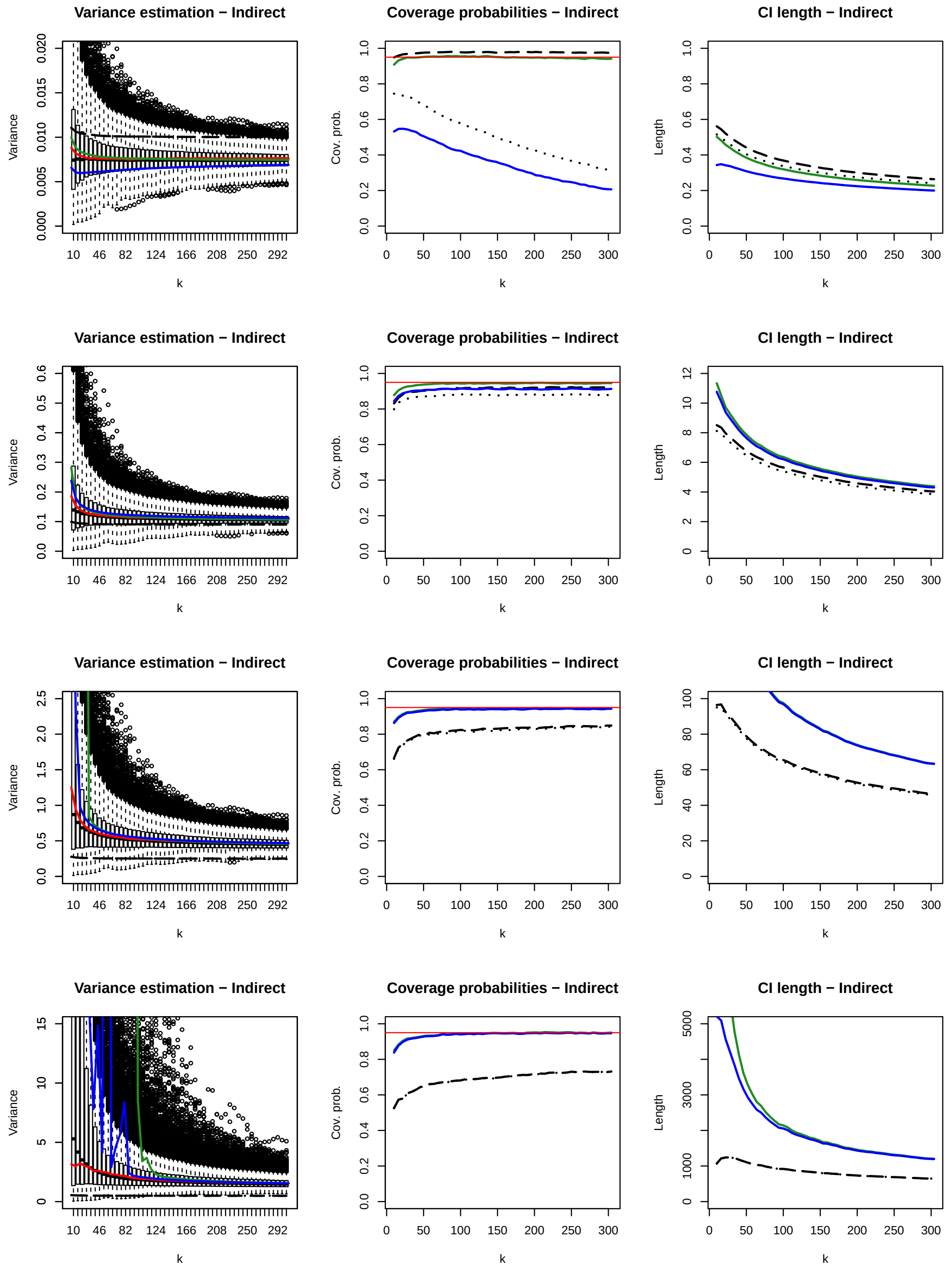

Fig. D6: As in Figure D4, with $n = 5{,}000$ and $\tau_n' = 1 - 1/n = 0.9998$.

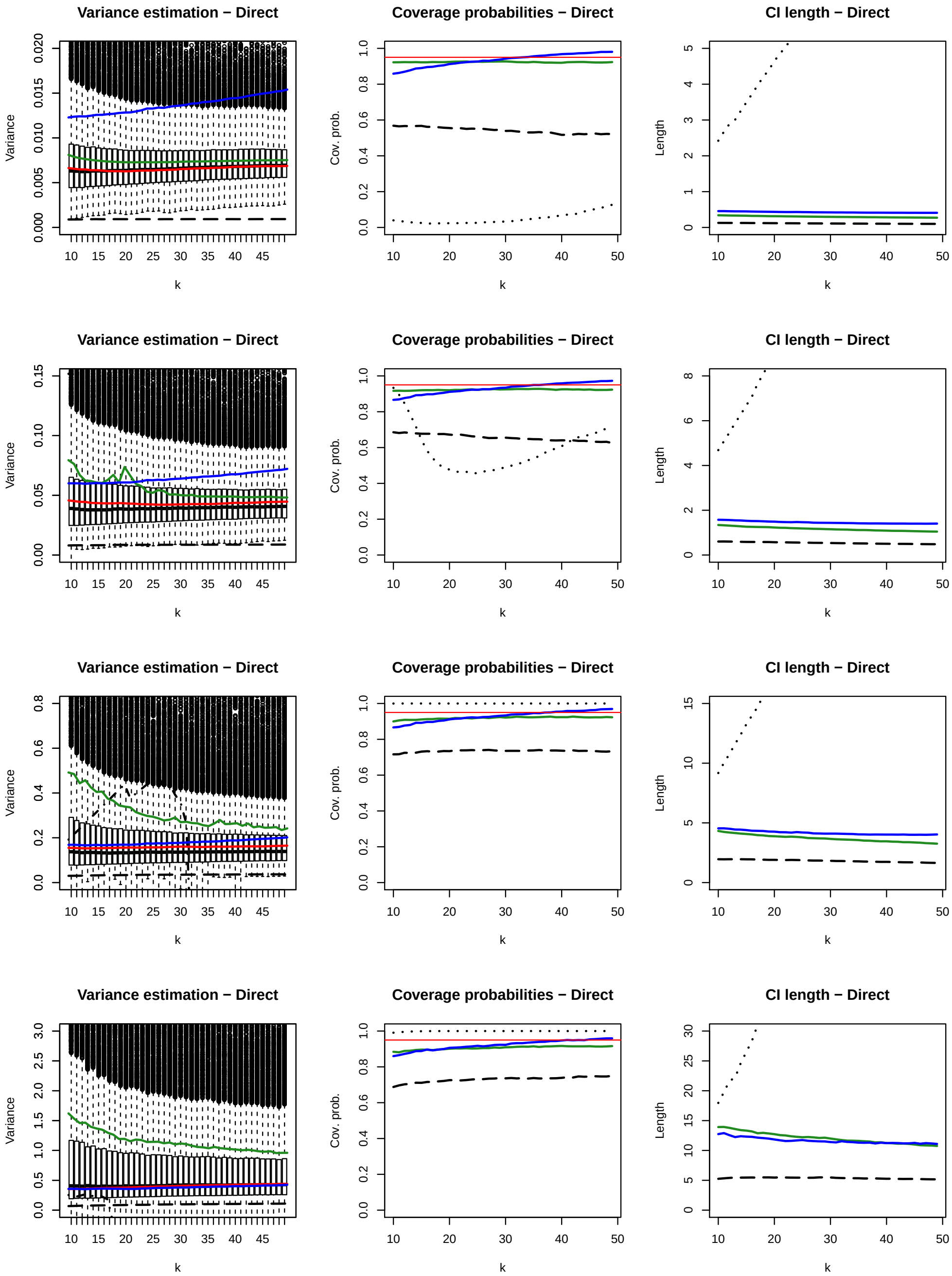

Fig. D7: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widehat{\xi}^{\star,\text{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widehat{s}^2_n(1)$ and the average values of $\widehat{s}^{2,\text{BR}}_n$ (dashed curve), $\widehat{s}^2_n(1)$ (green curve), $\widehat{s}^2_n$ (dashed-dotted curve) and $\overline{s}^2_n$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha;1)$, $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widehat{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Burr distribution with $\rho = -1$ and $\gamma = 0.1$ (first row), $\gamma = 0.2$ (second row), $\gamma = 0.3$ (third row), $\gamma = 0.4$ (fourth row). The sample size is $n = 200$ and the target expectile level is $\tau_n' = 1 - 1/n = 0.995$.

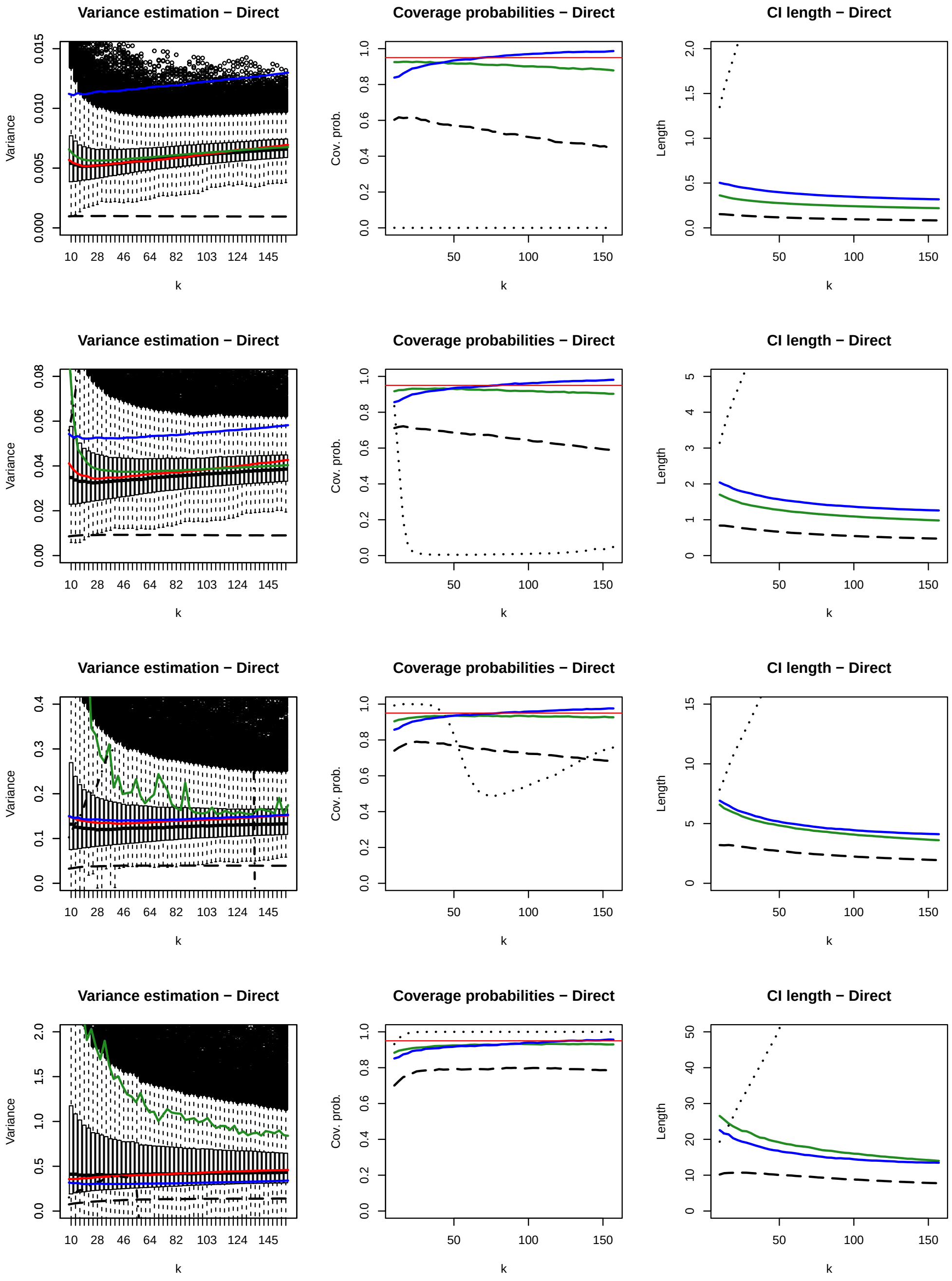

Fig. D8: As in Figure D7, with $n = 1{,}000$ and $\tau_n' = 1 - 1/n = 0.999$.

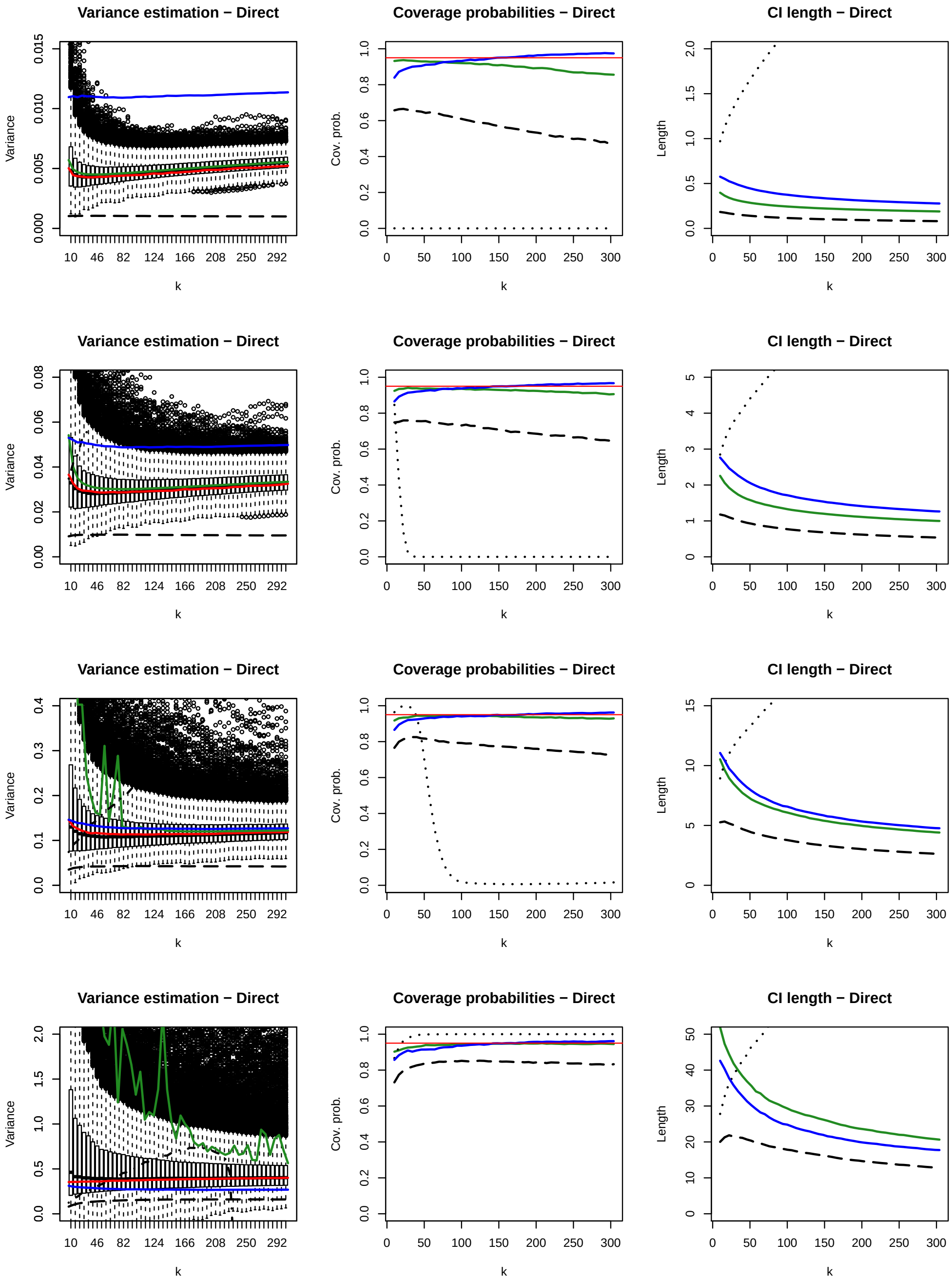

Fig. D9: As in Figure D7, with $n = 5{,}000$ and $\tau_n' = 1 - 1/n = 0.9998$.

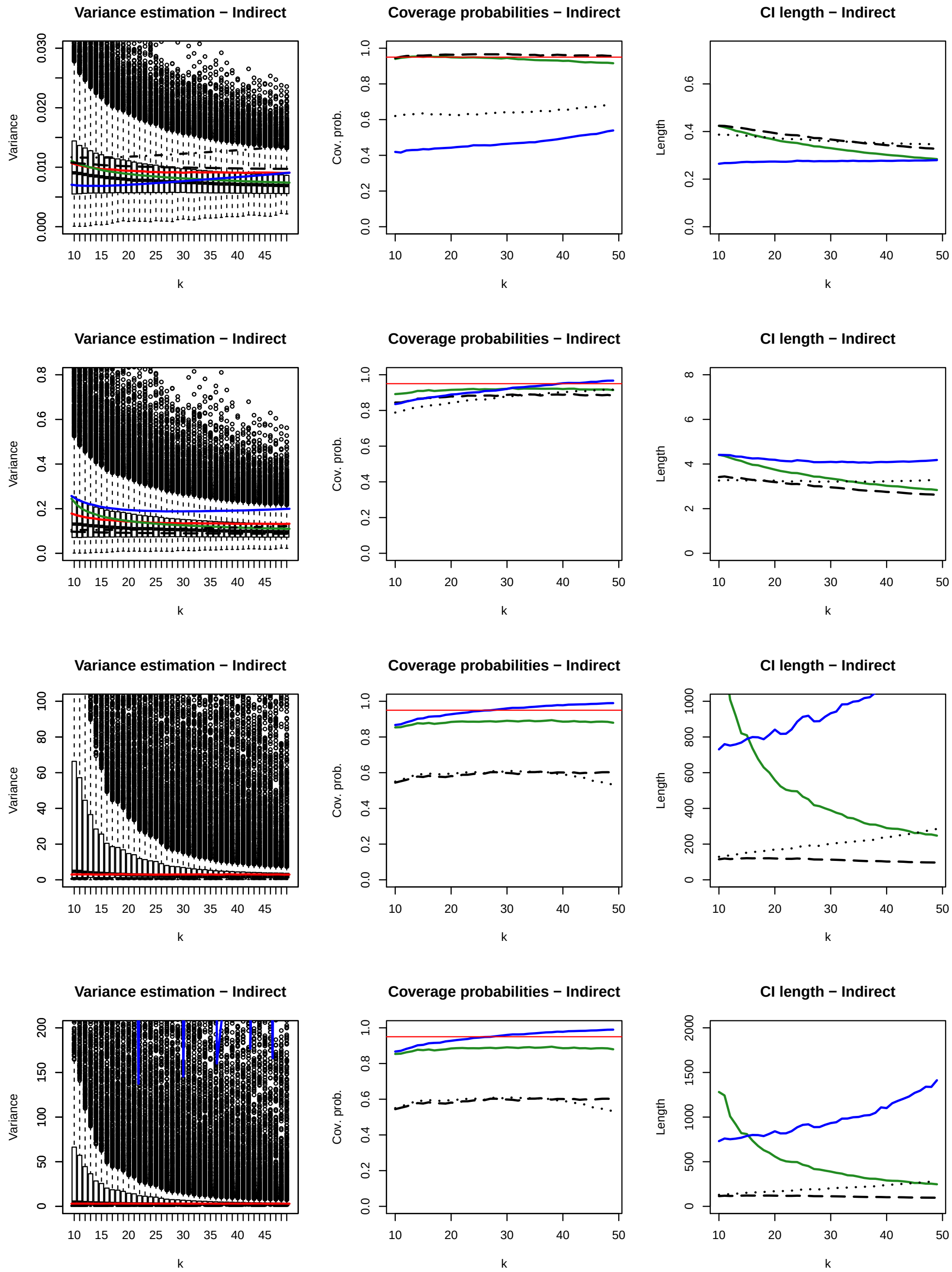

Fig. D10: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widetilde{\sigma}_n^2(2)$ and the average values of $\widetilde{\sigma}_n^{2,\mathrm{BR}}$ (dashed curve), $\widetilde{\sigma}_n^2(2)$ (green curve), $\widetilde{\sigma}_n^2$ (dashed-dotted curve) and $\overline{\sigma}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(2)}_{\tau_n'}(\alpha; 2)$, $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Burr distribution with $\rho = -1$ and $\gamma = 0.1$ (first row), $\gamma = 0.3$ (second row), $\gamma = 0.5$ (third row), $\gamma = 0.7$ (fourth row). The sample size is $n = 200$ and the target expectile level is $\tau_n' = 1 - 1/n = 0.995$.

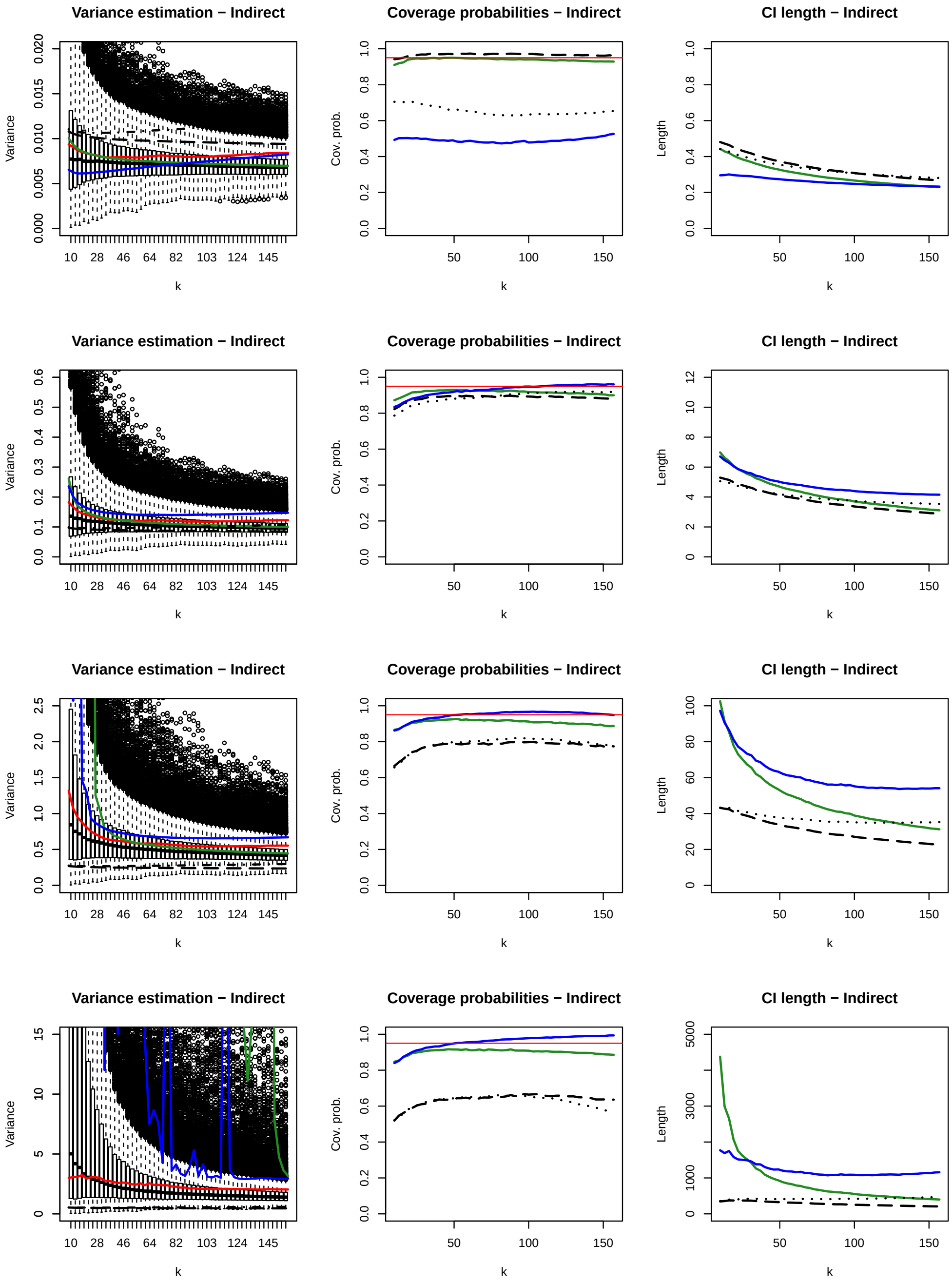

Fig. D11: As in Figure D10, with $n = 1{,}000$ and $\tau_n' = 1 - 1/n = 0.999$.

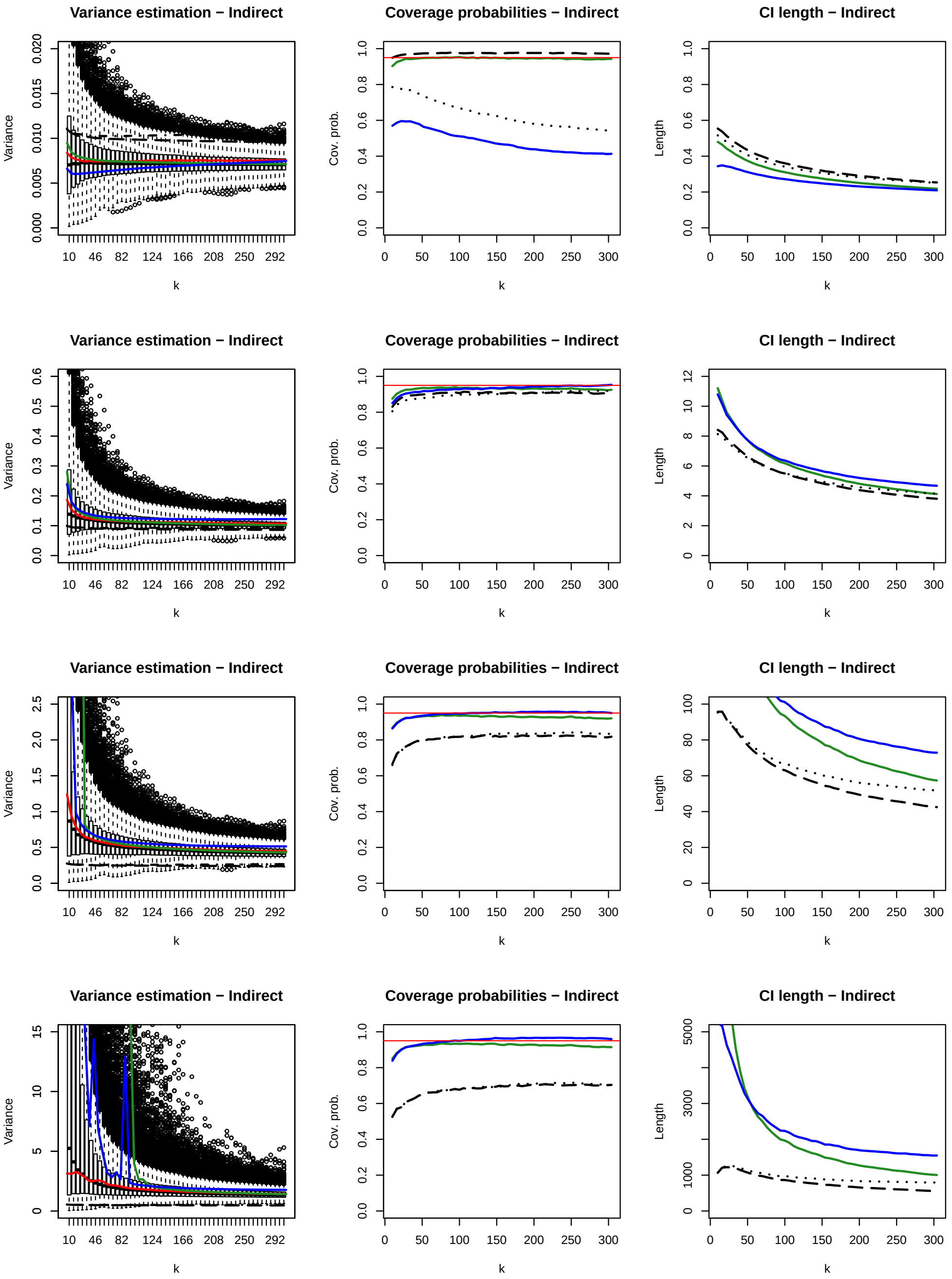

Fig. D12: As in Figure D10, with $n = 5{,}000$ and $\tau_n' = 1 - 1/n = 0.9998$.

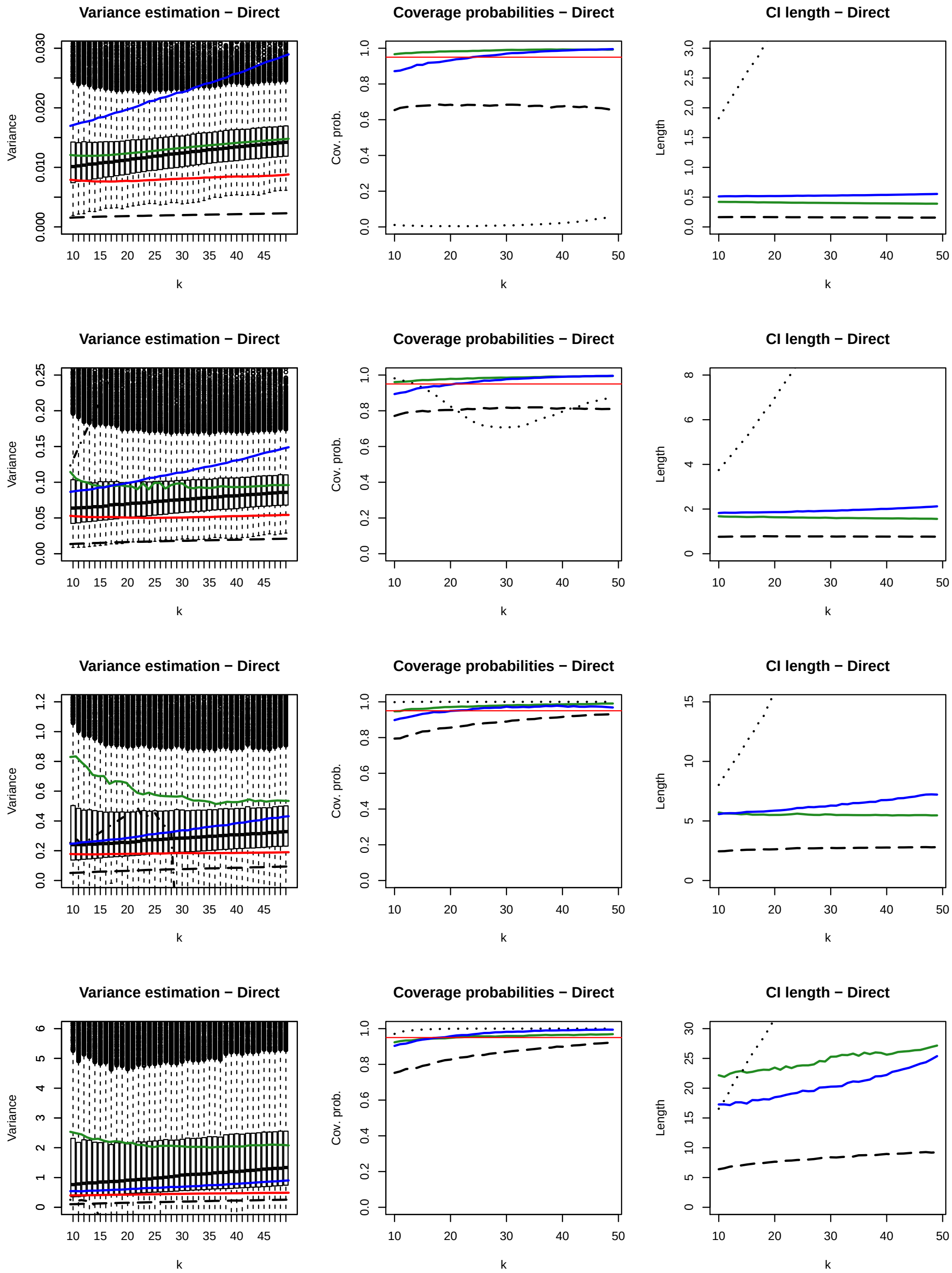

Fig. D13: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n')))\log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widehat{s}^2_n(1)$ and the average values of $\widehat{s}^{2,\mathrm{BR}}_n$ (dashed curve), $\widehat{s}^2_n(1)$ (green curve), $\widehat{s}^2_n$ (dashed-dotted curve) and $\overline{s}^2_n$ (blue curve) as functions of $k=k_n=n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha;1)$, $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widehat{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Burr distribution with $\rho=-1/2$ and $\gamma=0.1$ (first row), $\gamma=0.2$ (second row), $\gamma=0.3$ (third row), $\gamma=0.4$ (fourth row). The sample size is $n=200$ and the target expectile level is $\tau_n'=1-1/n=0.995$.

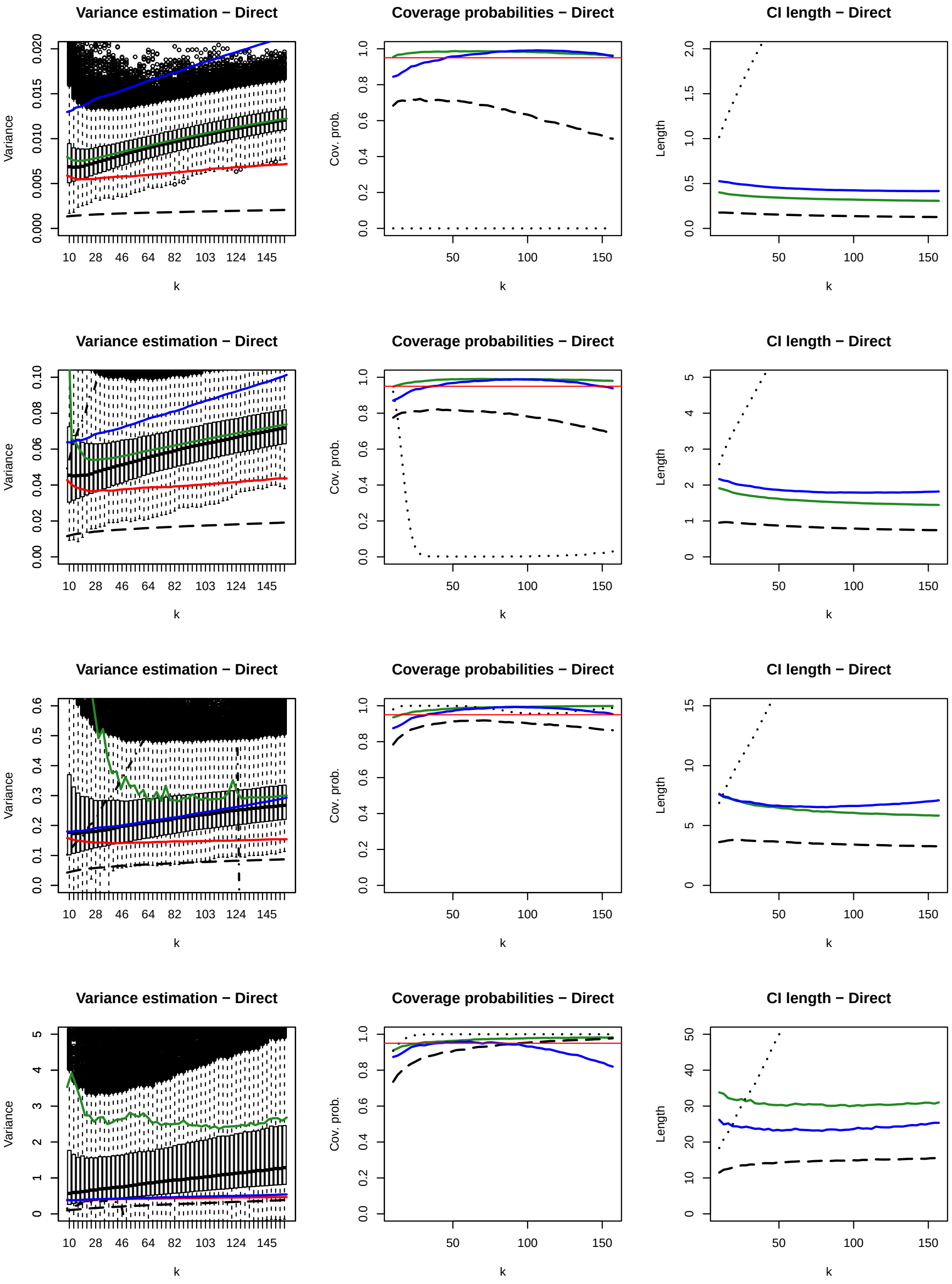

Fig. D14: As in Figure D13, with $n = 1{,}000$ and $\tau_n' = 1 - 1/n = 0.999$.

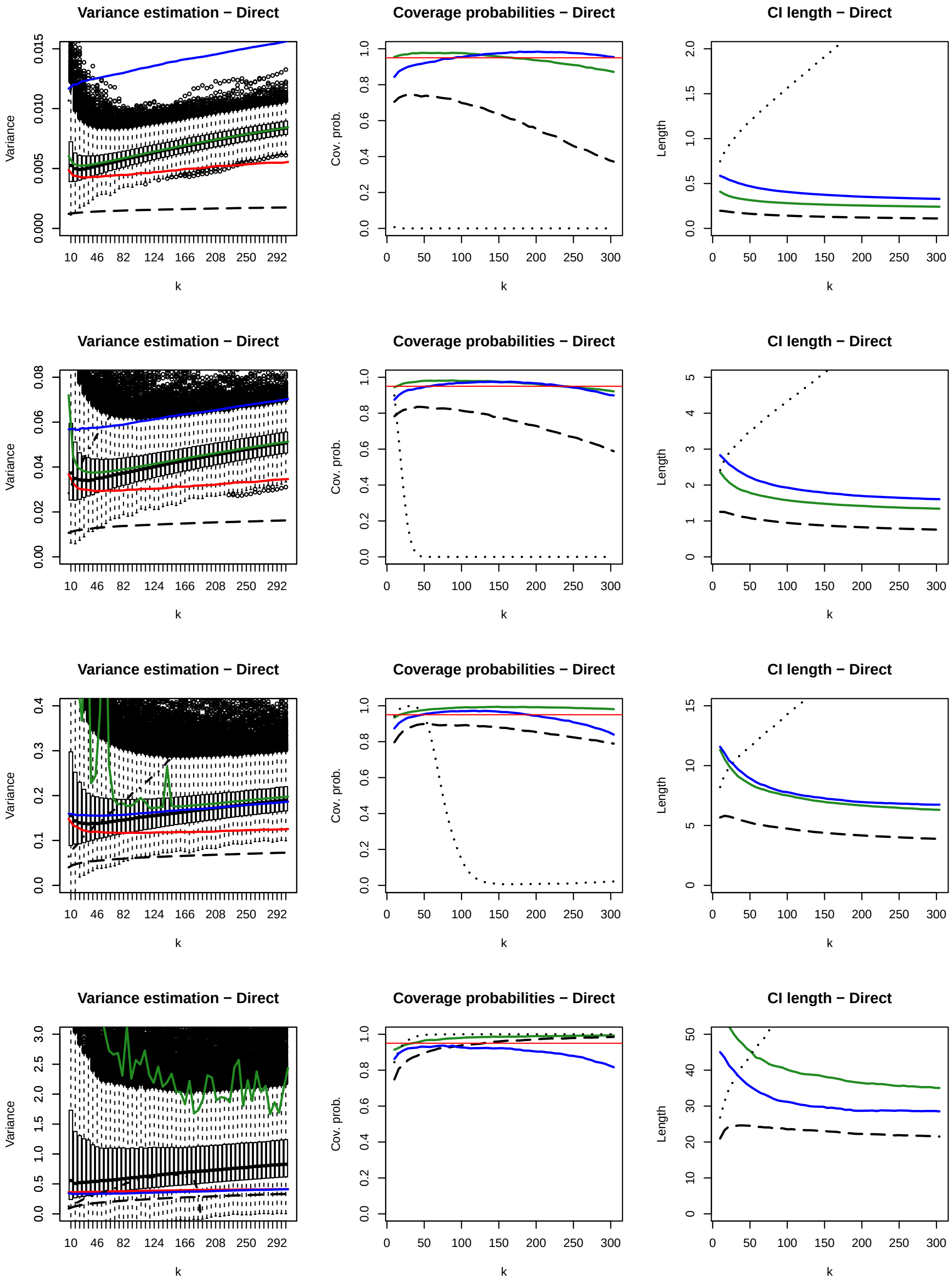

Fig. D15: As in Figure D13, with $n = 5{,}000$ and $\tau_n' = 1 - 1/n = 0.9998$.

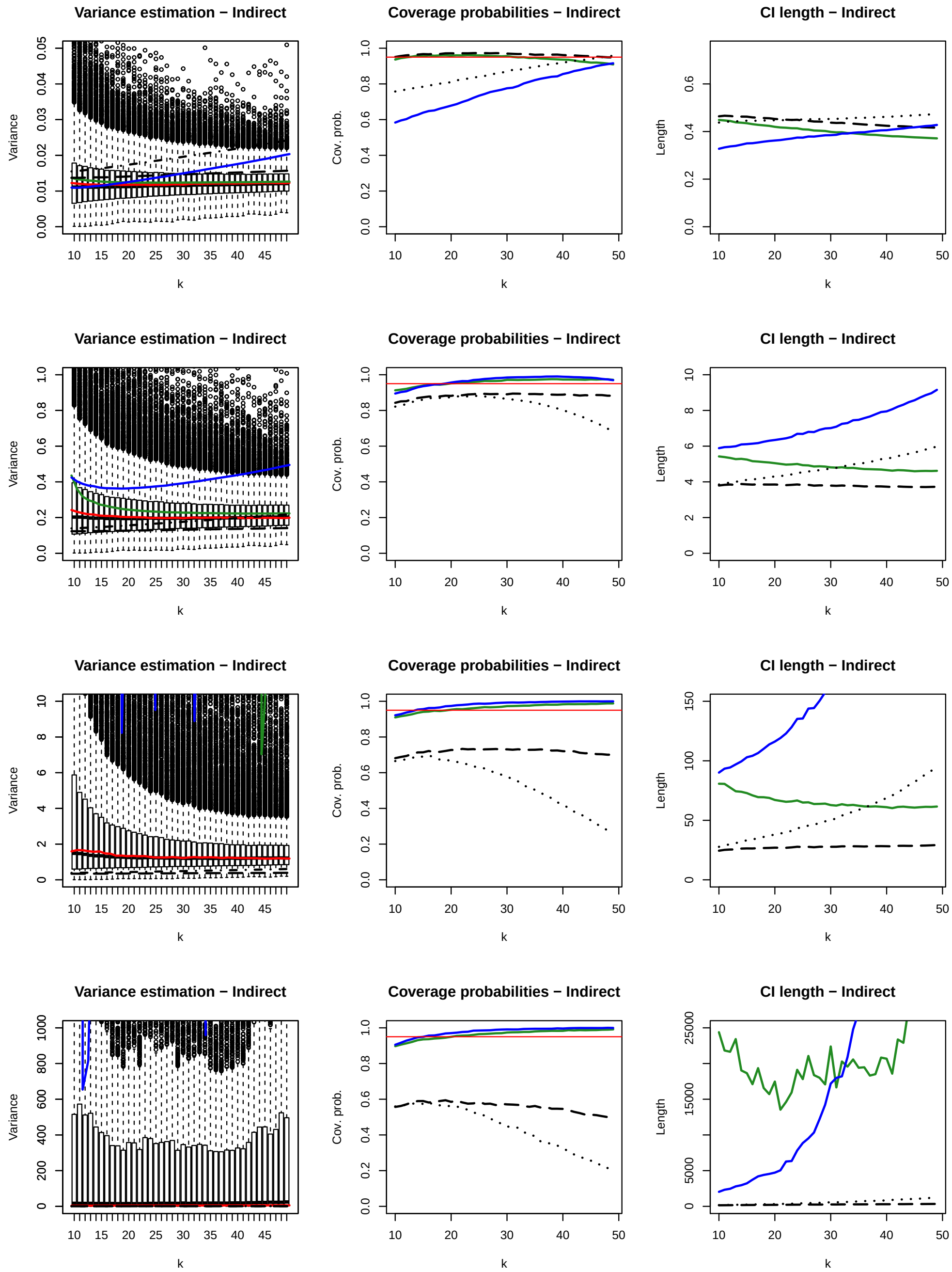

Fig. D16: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widetilde{\sigma}_n^2(2)$ and the average values of $\widetilde{\sigma}_n^{2,\mathrm{BR}}$ (dashed curve), $\widetilde{\sigma}_n^2(2)$ (green curve), $\widetilde{\sigma}_n^2$ (dashed-dotted curve) and $\overline{\sigma}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(2)}_{\tau_n'}(\alpha;2)$, $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Burr distribution with $\rho = -0.5$ and $\gamma = 0.1$ (first row), $\gamma = 0.3$ (second row), $\gamma = 0.5$ (third row), $\gamma = 0.7$ (fourth row). The sample size is $n = 200$ and the target expectile level is $\tau_n' = 1 - 1/n = 0.995$.

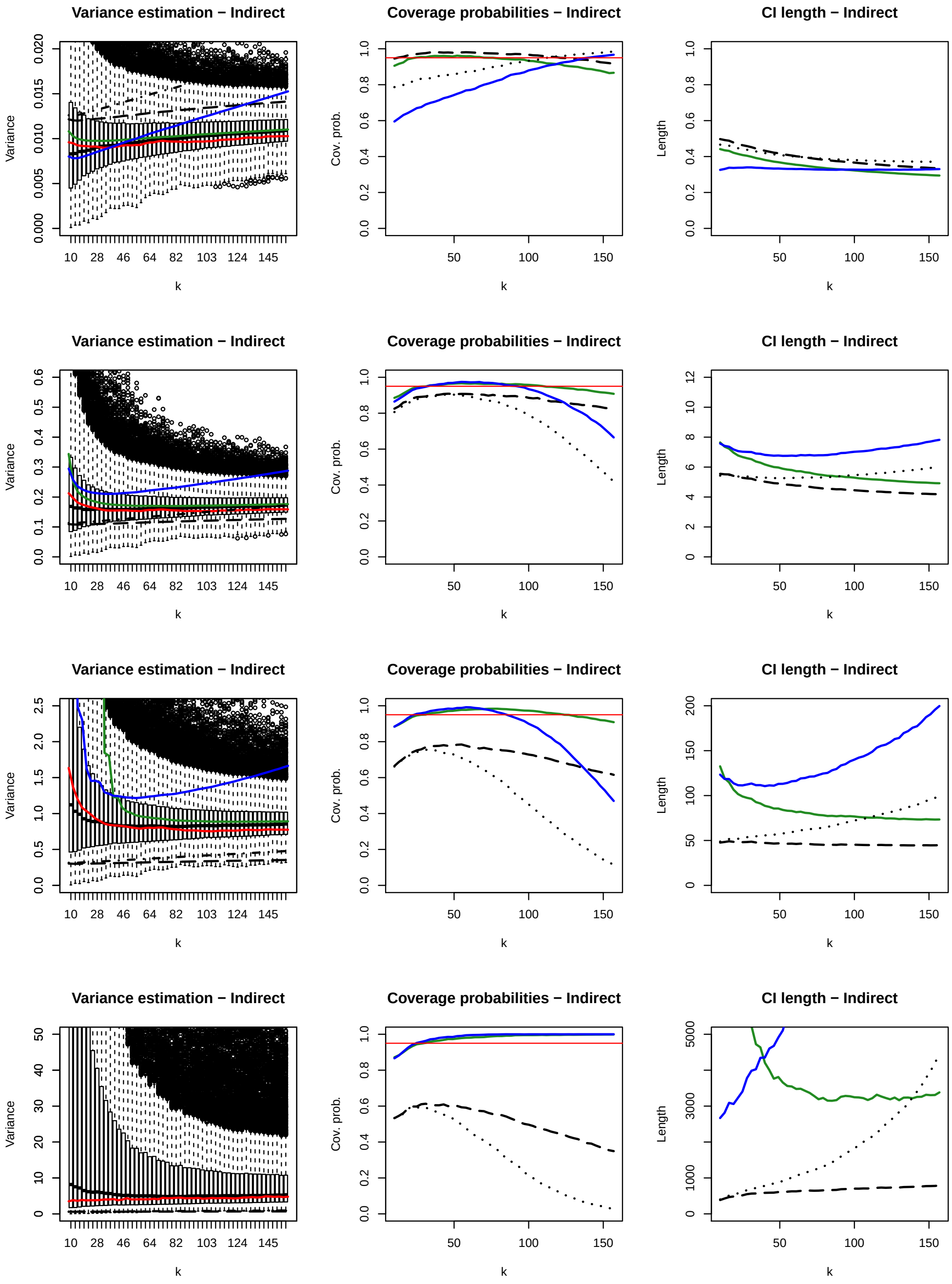

Fig. D17: As in Figure D16, with $n = 1{,}000$ and $\tau'_n = 1 - 1/n = 0.999$.

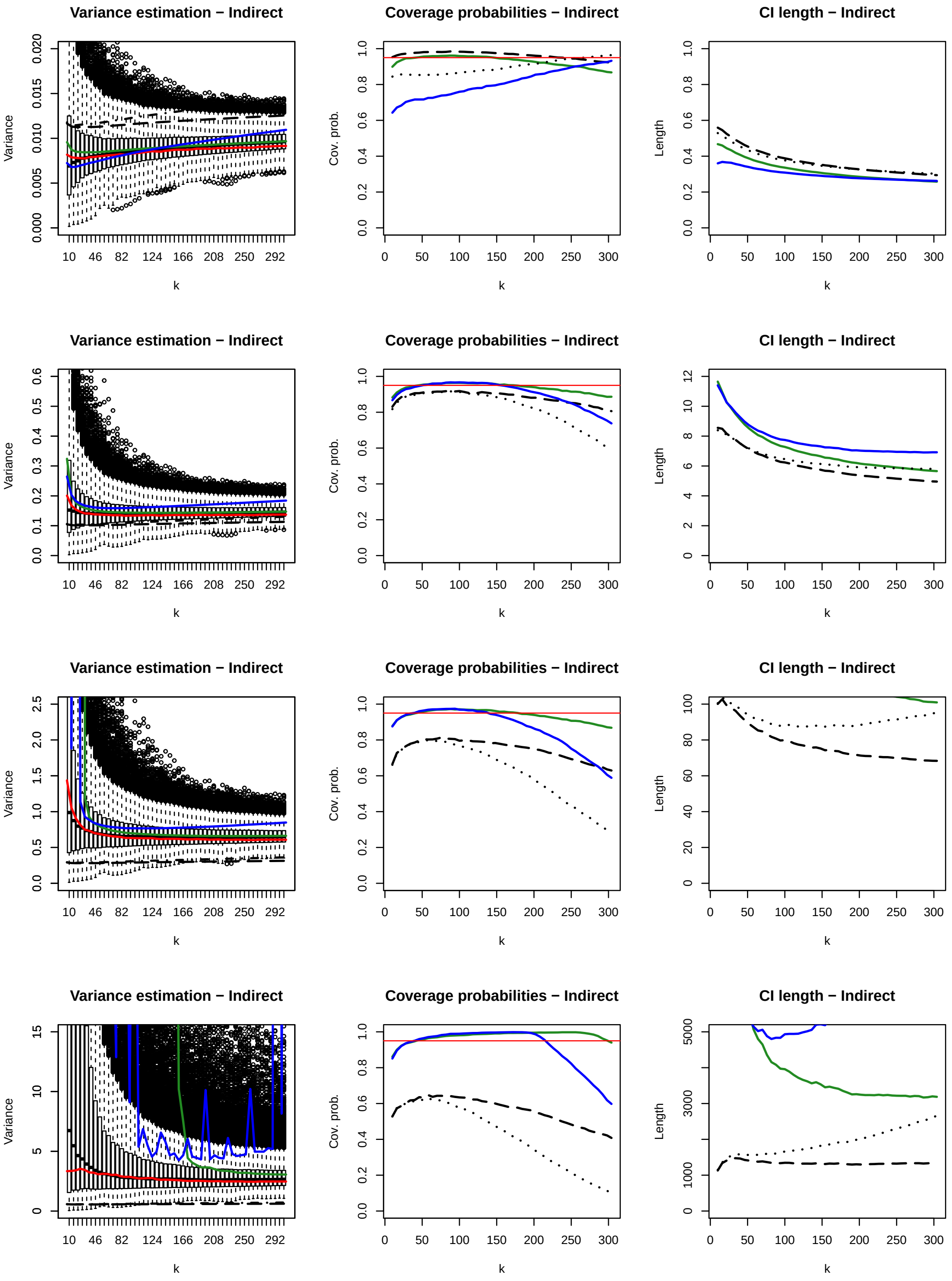

Fig. D18: As in Figure D16, with $n = 5{,}000$ and $\tau_n' = 1 - 1/n = 0.9998$.

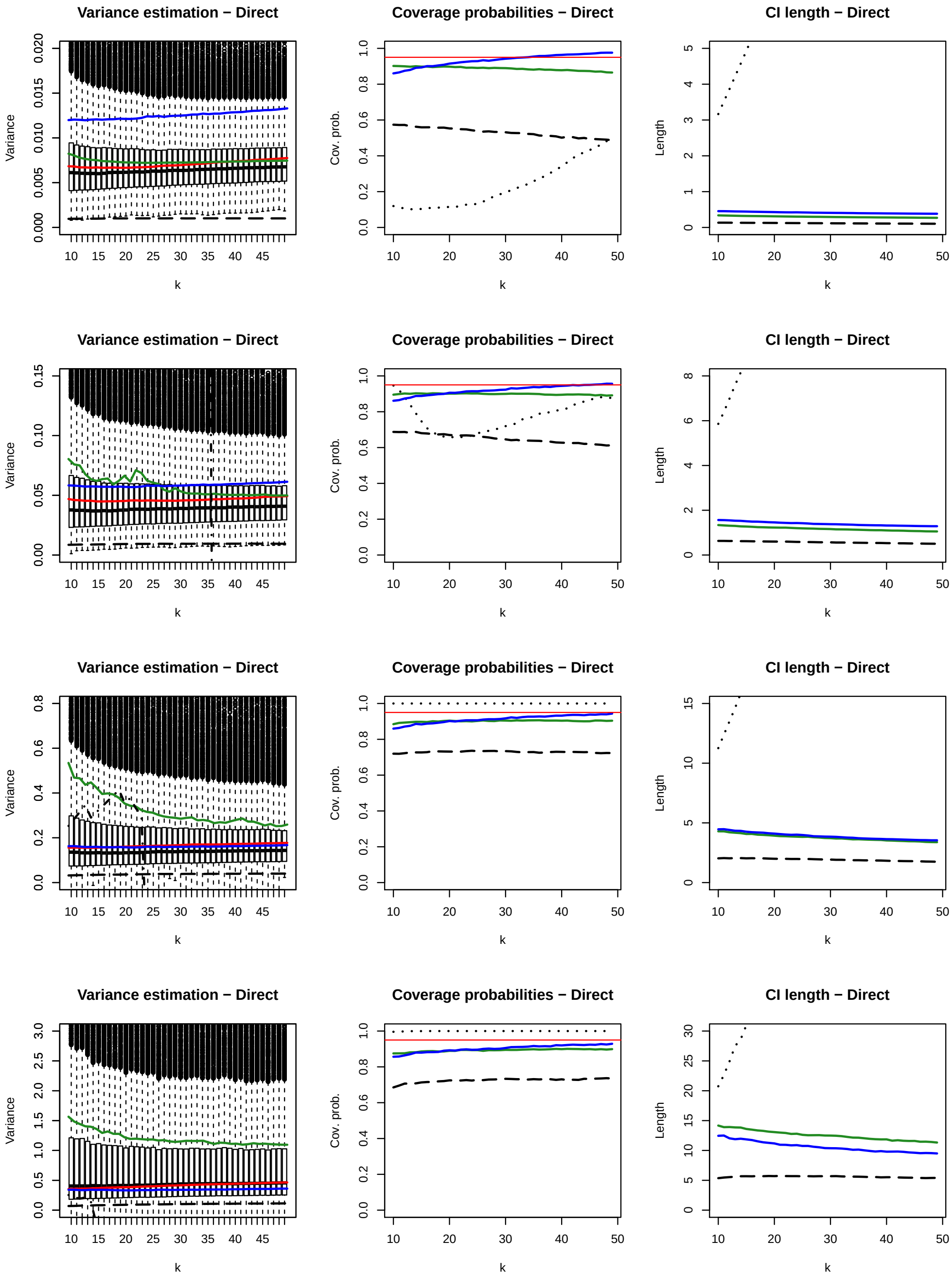

Fig. D19: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widehat{s}_n^2(1)$ and the average values of $\widehat{s}_n^{2,\mathrm{BR}}$ (dashed curve), $\widehat{s}_n^2(1)$ (green curve), $\widehat{s}_n^2$ (dashed-dotted curve) and $\overline{s}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha; 1)$, $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widehat{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Fréchet distribution with $\gamma = 0.1$ (first row), $\gamma = 0.2$ (second row), $\gamma = 0.3$ (third row), $\gamma = 0.4$ (fourth row). The sample size is $n = 200$ and the target expectile level is $\tau_n' = 1 - 1/n = 0.995$.

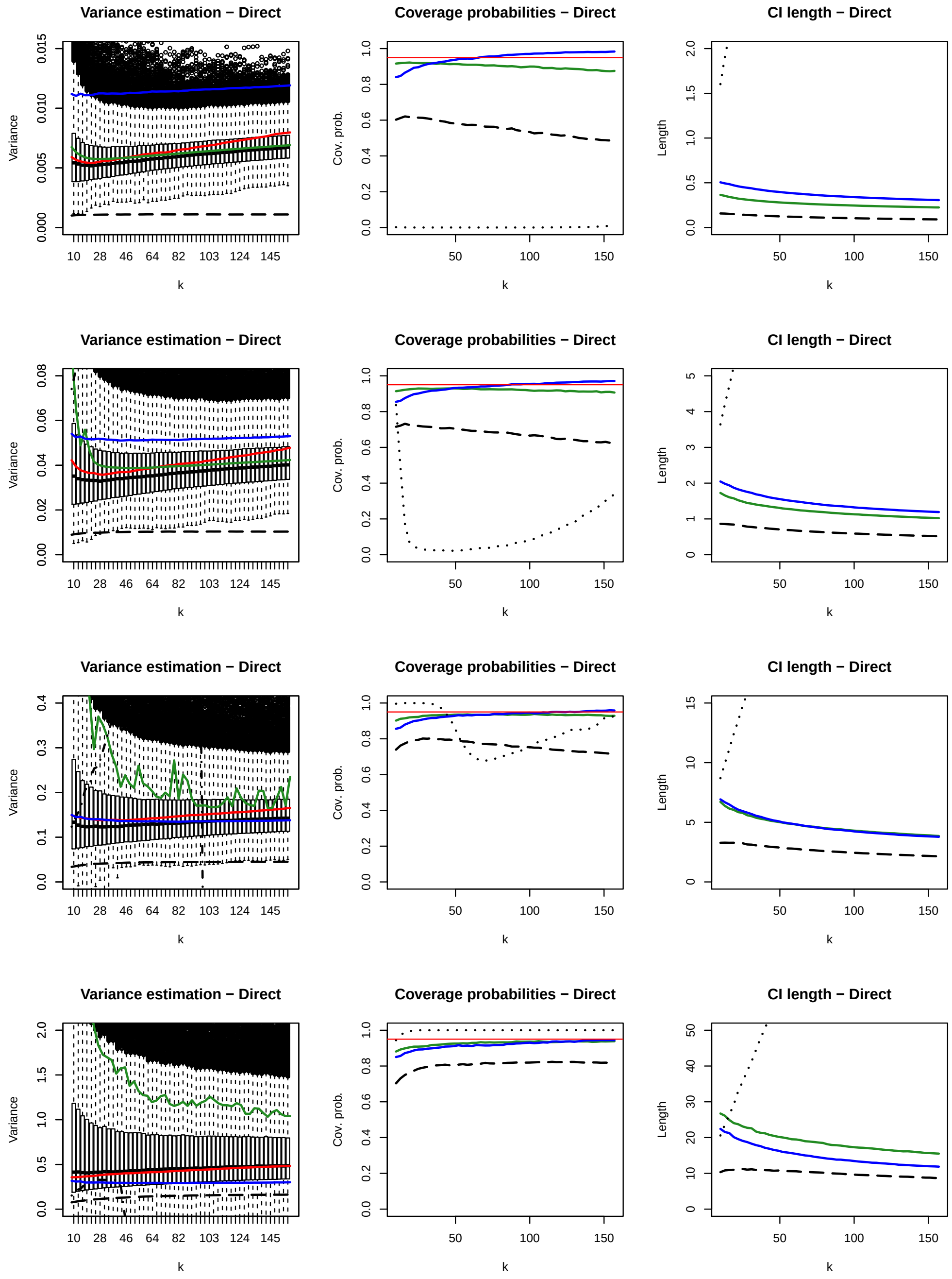

Fig. D20: As in Figure D19, with $n = 1{,}000$ and $\tau_n' = 1 - 1/n = 0.999$.

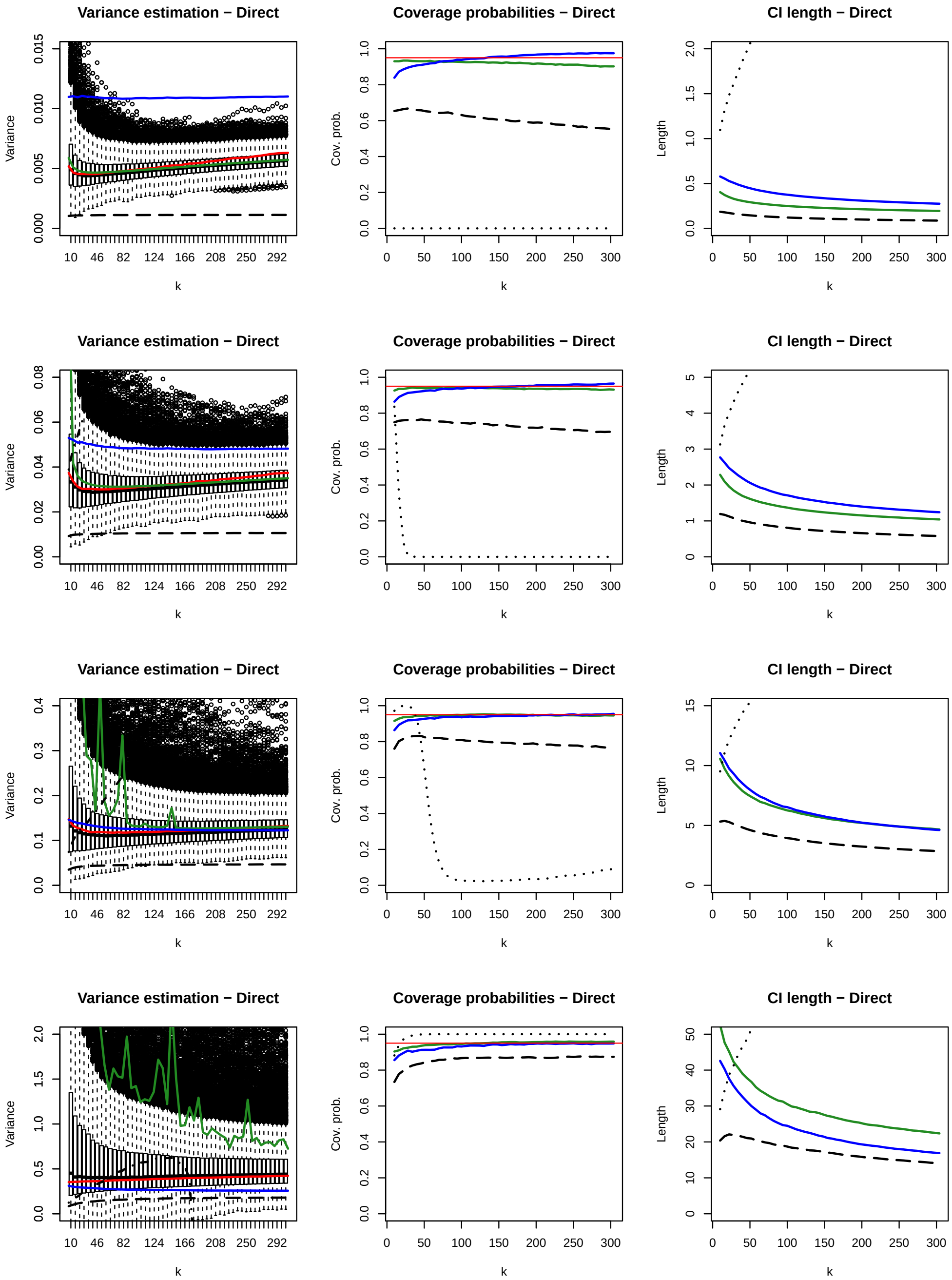

Fig. D21: As in Figure D19, with $n = 5{,}000$ and $\tau_n' = 1 - 1/n = 0.9998$.

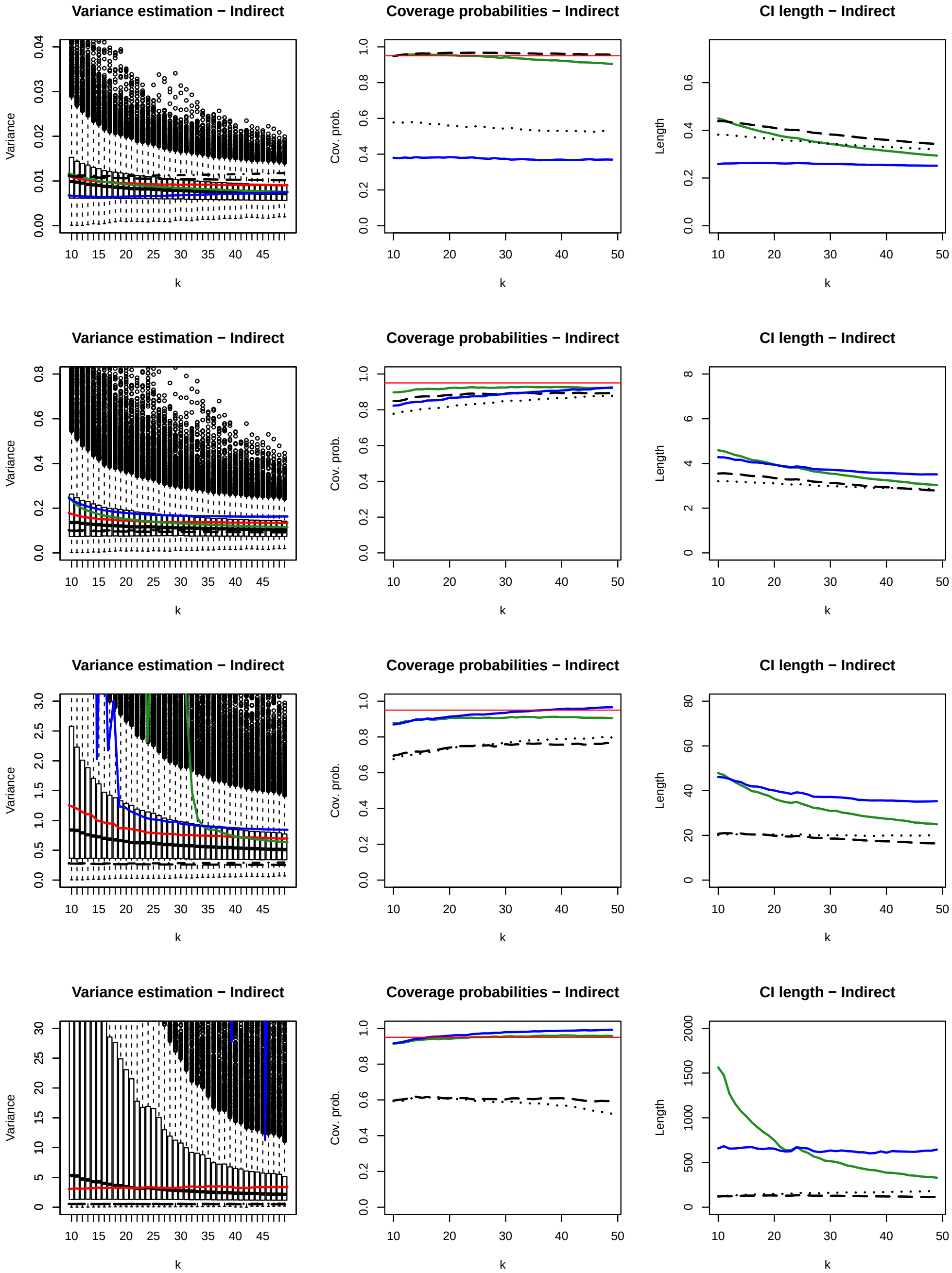

Fig. D22: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widetilde{\sigma}_n^2(2)$ and the average values of $\widetilde{\sigma}_n^{2,\mathrm{BR}}$ (dashed curve), $\widetilde{\sigma}_n^2(2)$ (green curve), $\widetilde{\sigma}_n^2$ (dashed-dotted curve) and $\overline{\sigma}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(2)}_{\tau_n'}(\alpha; 2)$, $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Fréchet distribution with $\gamma = 0.1$ (first row), $\gamma = 0.3$ (second row), $\gamma = 0.5$ (third row), $\gamma = 0.7$ (fourth row). The sample size is $n = 200$ and the target expectile level is $\tau_n' = 1 - 1/n = 0.995$.

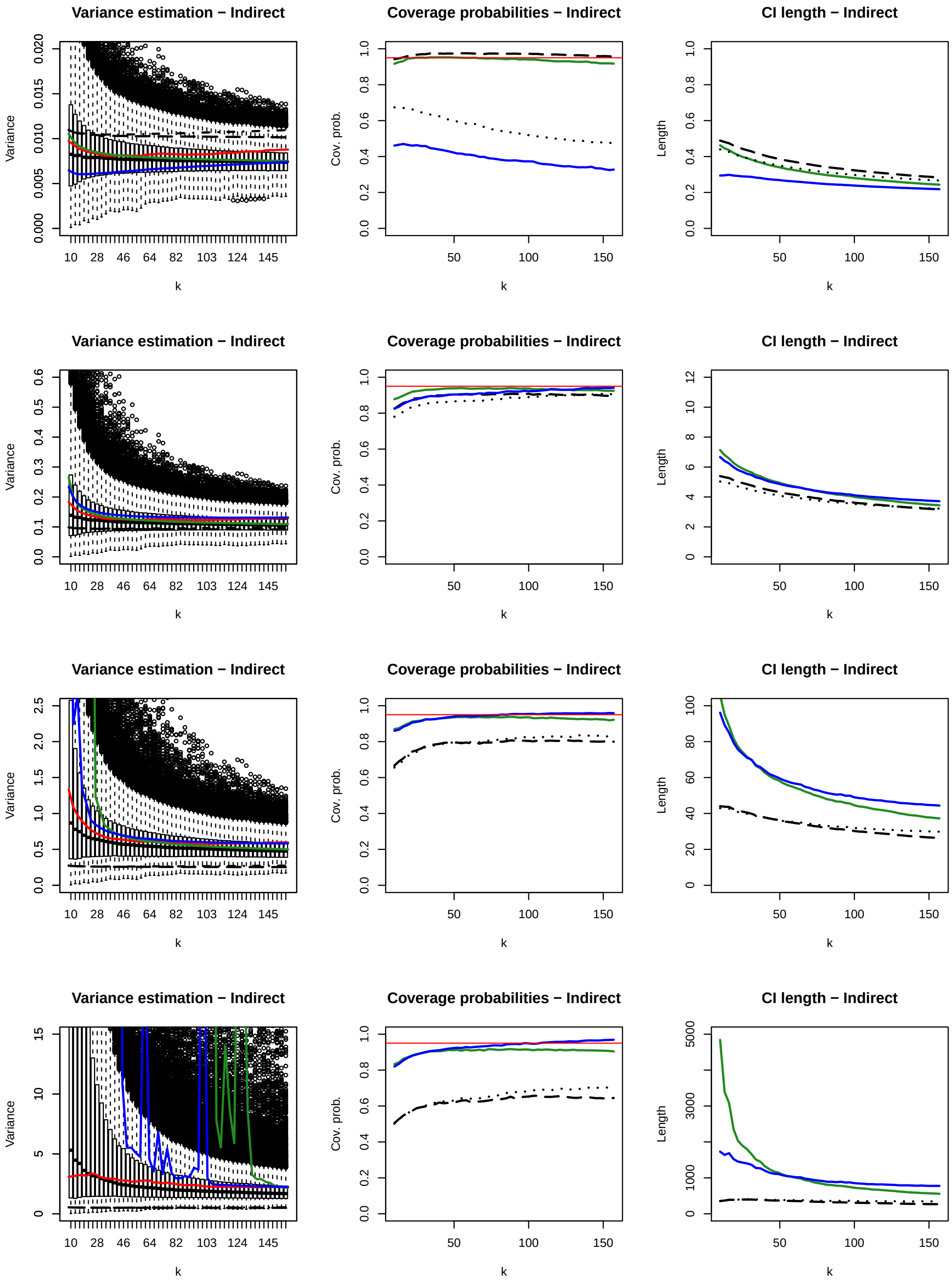

Fig. D23: As in Figure D22, with $n = 1{,}000$ and $\tau_n' = 1 - 1/n = 0.999$.

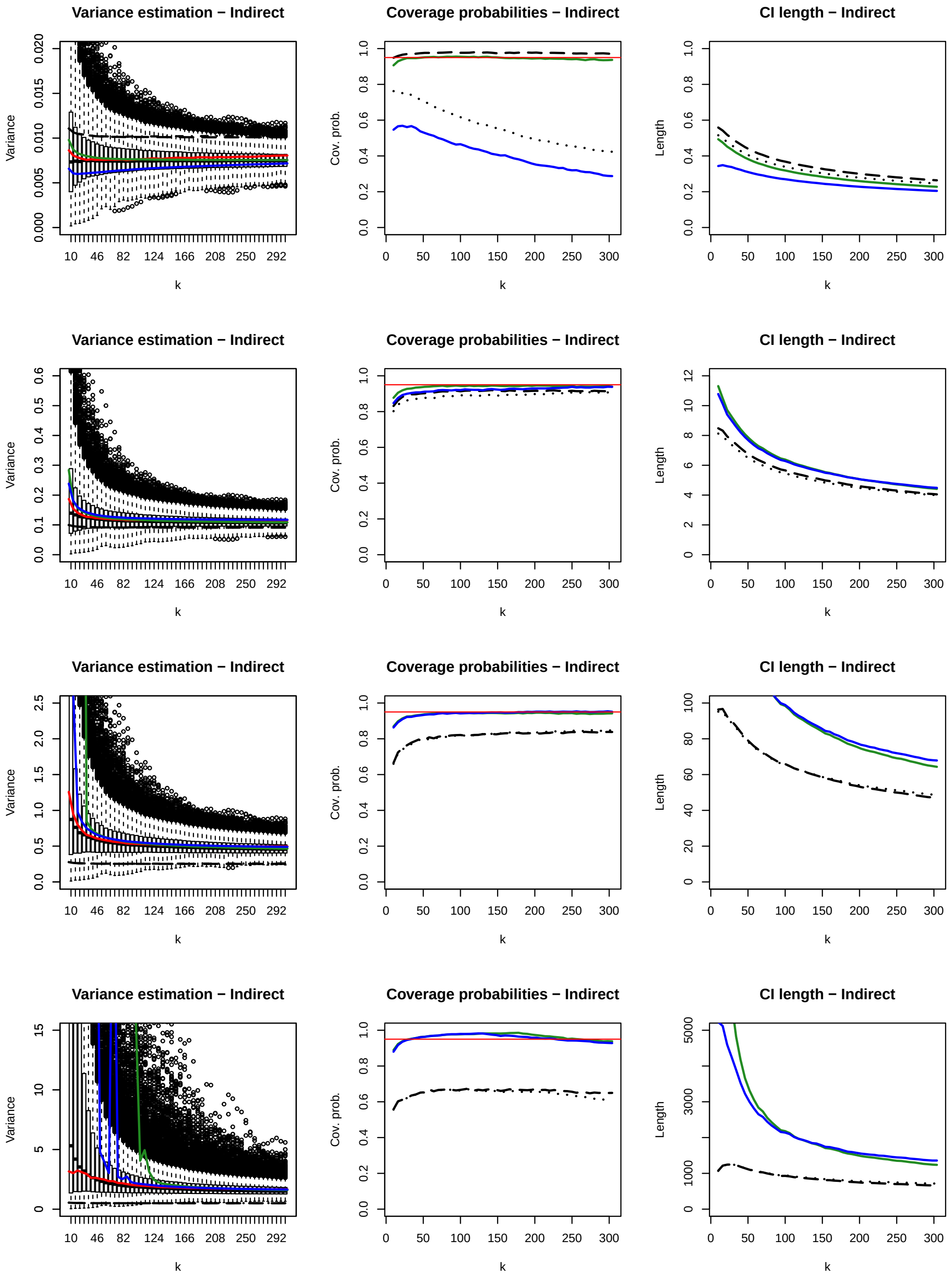

Fig. D24: As in Figure D22, with $n = 5{,}000$ and $\tau_n' = 1 - 1/n = 0.9998$.

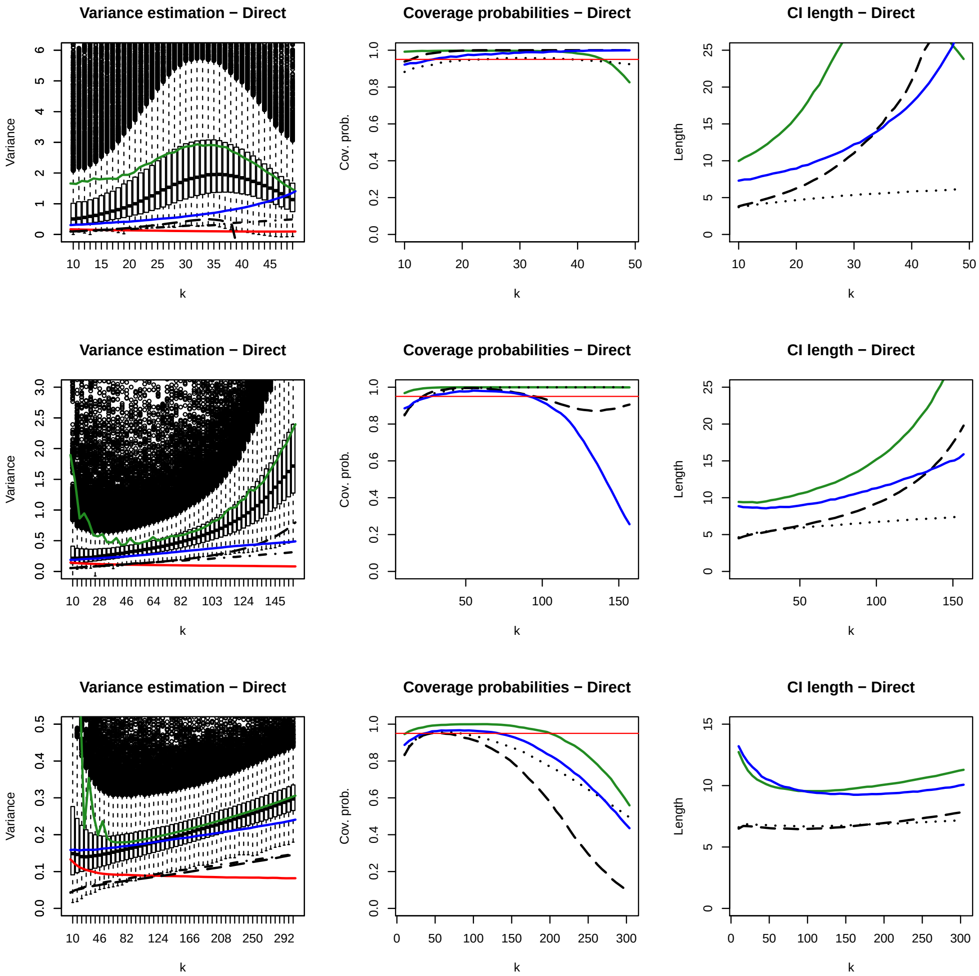

Fig. D25: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n')))\log(\widehat{\xi}^{\star,\text{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widehat{s}_n^2(1)$ and the average values of $\widehat{s}_n^{2,\text{BR}}$ (dashed curve), $\widehat{s}_n^2(1)$ (green curve), $\widehat{s}_n^2$ (dashed-dotted curve) and $\overline{s}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha;1)$, $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widehat{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Student distribution with $\nu = 10/3$ degrees of freedom, resulting in $\gamma = 0.3$. The sample size is $n = 200$ (first row), 1,000 (second row) and 5,000 (third row) and the target expectile level is $\tau_n' = 1 - 1/n$.

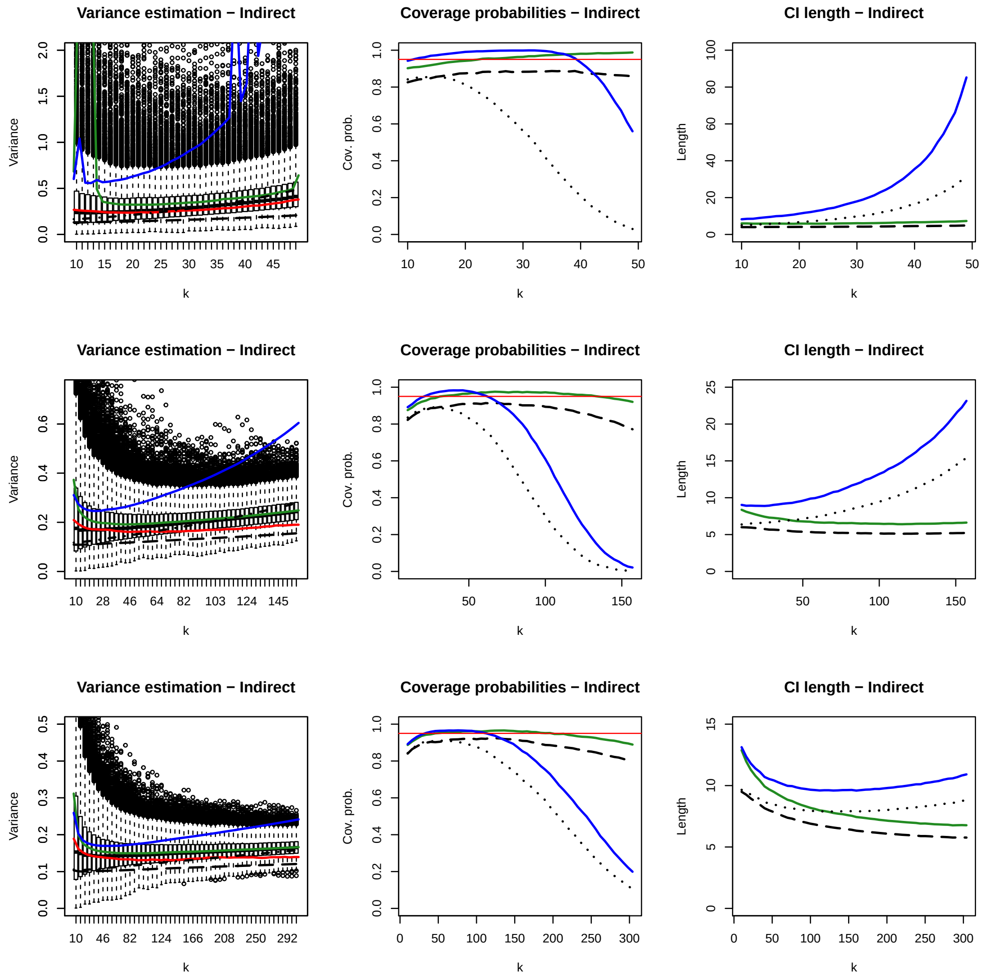

Fig. D26: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widetilde{\sigma}_n^2(2)$ and the average values of $\widetilde{\sigma}_n^{2,\mathrm{BR}}$ (dashed curve), $\widetilde{\sigma}_n^2(2)$ (green curve), $\widetilde{\sigma}_n^2$ (dashed-dotted curve) and $\overline{\sigma}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(2)}_{\tau_n'}(\alpha;2)$, $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Student distribution with $\nu = 10/3$ degrees of freedom, resulting in $\gamma = 0.3$. The sample size is $n = 200$ (first row), 1,000 (second row) and 5,000 (third row) and the target expectile level is $\tau_n' = 1 - 1/n$.

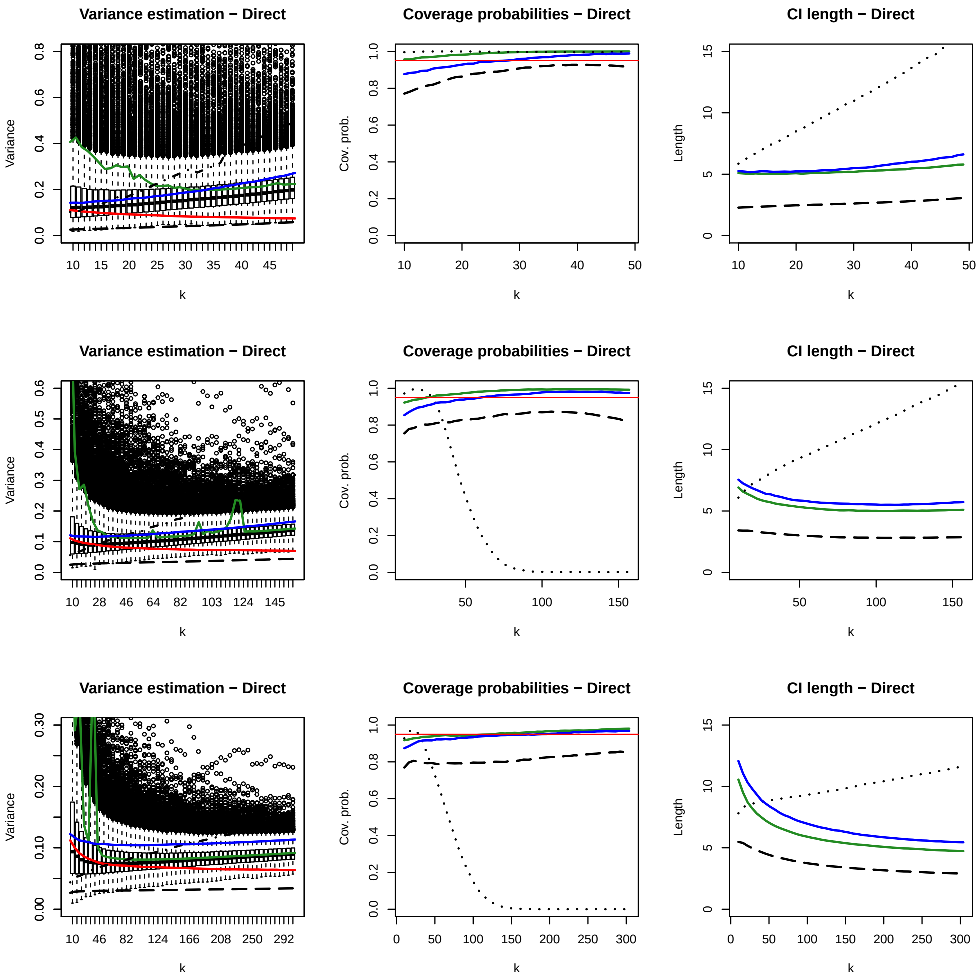

Fig. D27: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n')))\log(\widehat{\xi}^{\star,\text{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widehat{s}_n^2(1)$ and the average values of $\widehat{s}_n^{2,\text{BR}}$ (dashed curve), $\widehat{s}_n^2(1)$ (green curve), $\widehat{s}_n^2$ (dashed-dotted curve) and $\overline{s}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha;1)$, $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widehat{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Student distribution with $\nu = 10/3$ degrees of freedom and location parameter 1 (*i.e.* distributed as $Z+1$, where Z is Student distributed with $\nu = 10/3$ degrees of freedom), resulting in $\gamma = 0.3$. The sample size is $n = 200$ (first row), 1,000 (second row) and 5,000 (third row) and the target expectile level is $\tau_n' = 1 - 1/n$.

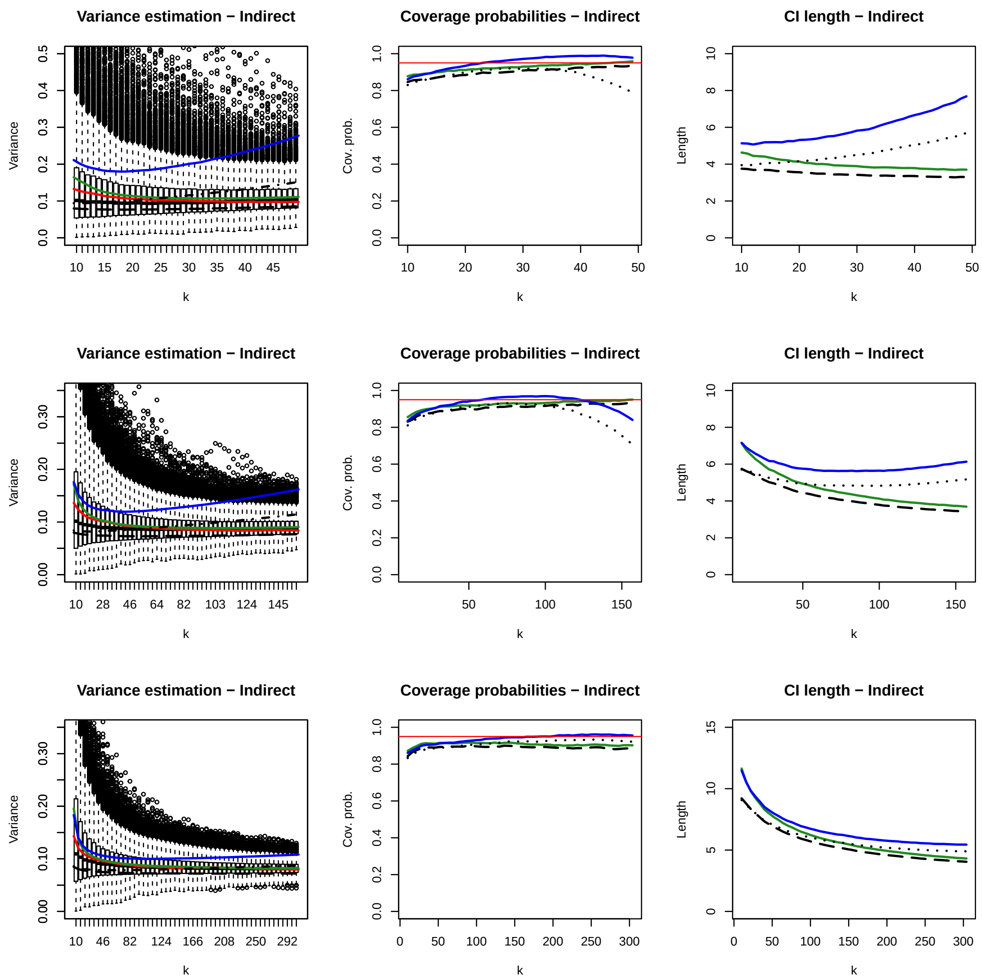

Fig. D28: Left panels: Comparison of the empirical variance of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))) \log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ (red curve) with boxplots of the asymptotic variance estimate $\widetilde{\sigma}_n^2(2)$ and the average values of $\widetilde{\sigma}_n^{2,\mathrm{BR}}$ (dashed curve), $\widetilde{\sigma}_n^2(2)$ (green curve), $\widetilde{\sigma}_n^2$ (dashed-dotted curve) and $\overline{\sigma}_n^2$ (blue curve) as functions of $k = k_n = n(1-\tau_n)$. Middle panels: coverage probabilities of the associated 95% confidence intervals $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(2)}_{\tau_n'}(\alpha; 2)$, $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(4)}_{\tau_n'}(\alpha)$ (respectively in black dashed line, green line, black dotted line and blue line) with the nominal level 0.95 in red. Right panels: Median length of these confidence intervals (same color code as in the middle panels). The underlying distribution is a Student distribution with $\nu = 10/3$ degrees of freedom and location parameter 1 (*i.e.* distributed as $Z+1$, where Z is Student distributed with $\nu = 10/3$ degrees of freedom), resulting in $\gamma = 0.3$. The sample size is $n = 200$ (first row), 1,000 (second row) and 5,000 (third row) and the target expectile level is $\tau_n' = 1 - 1/n$.

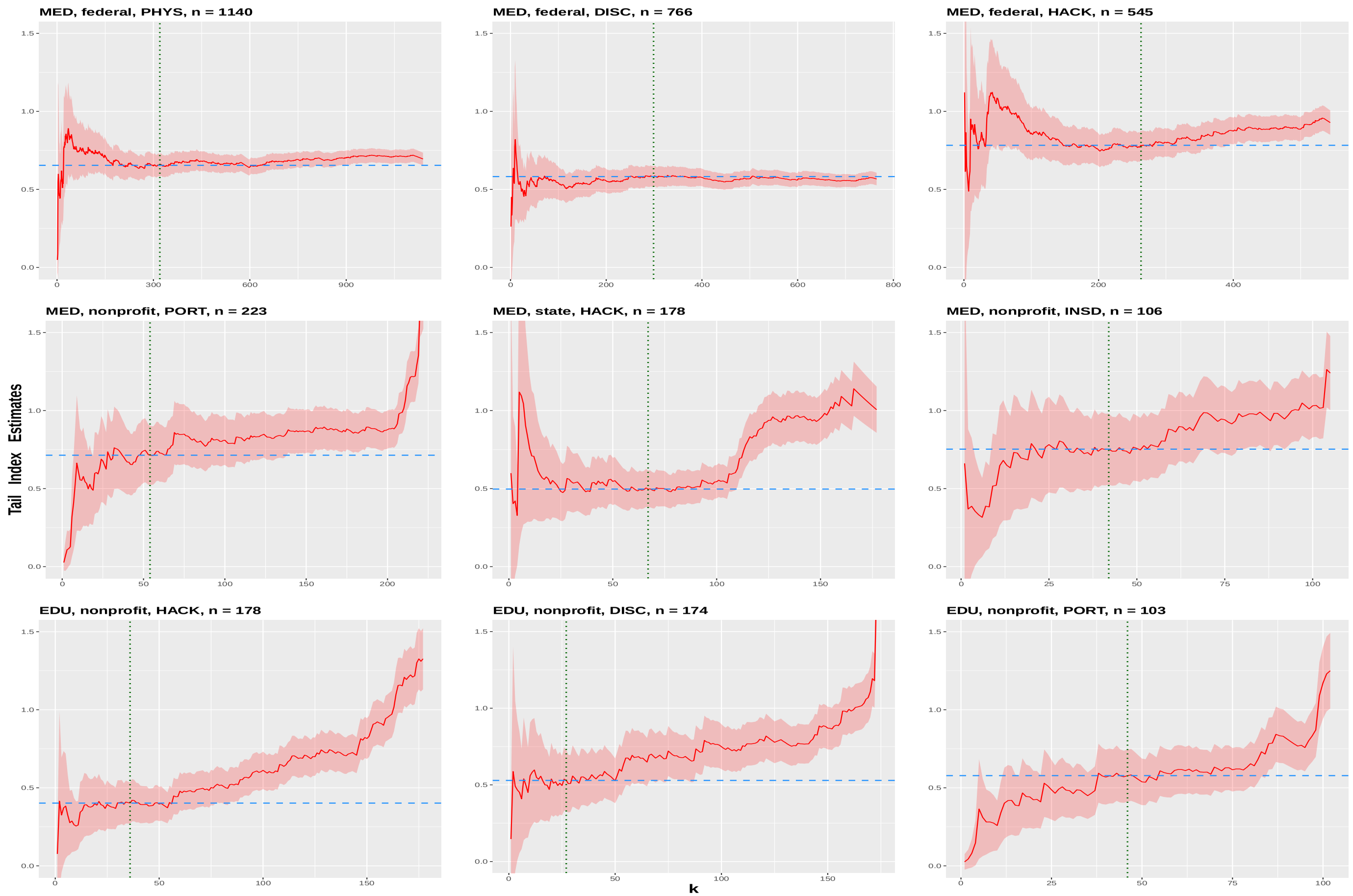

Fig. D29: For each cluster in Table 2, the corresponding plot of the bias-reduced Hill estimator versus k (red line), along with the asymptotic Gaussian 95% confidence interval (lighter red bands), the value of k selected by the path stability procedure (vertical dotted line) and its associated pointwise estimate (horizontal dashed line).

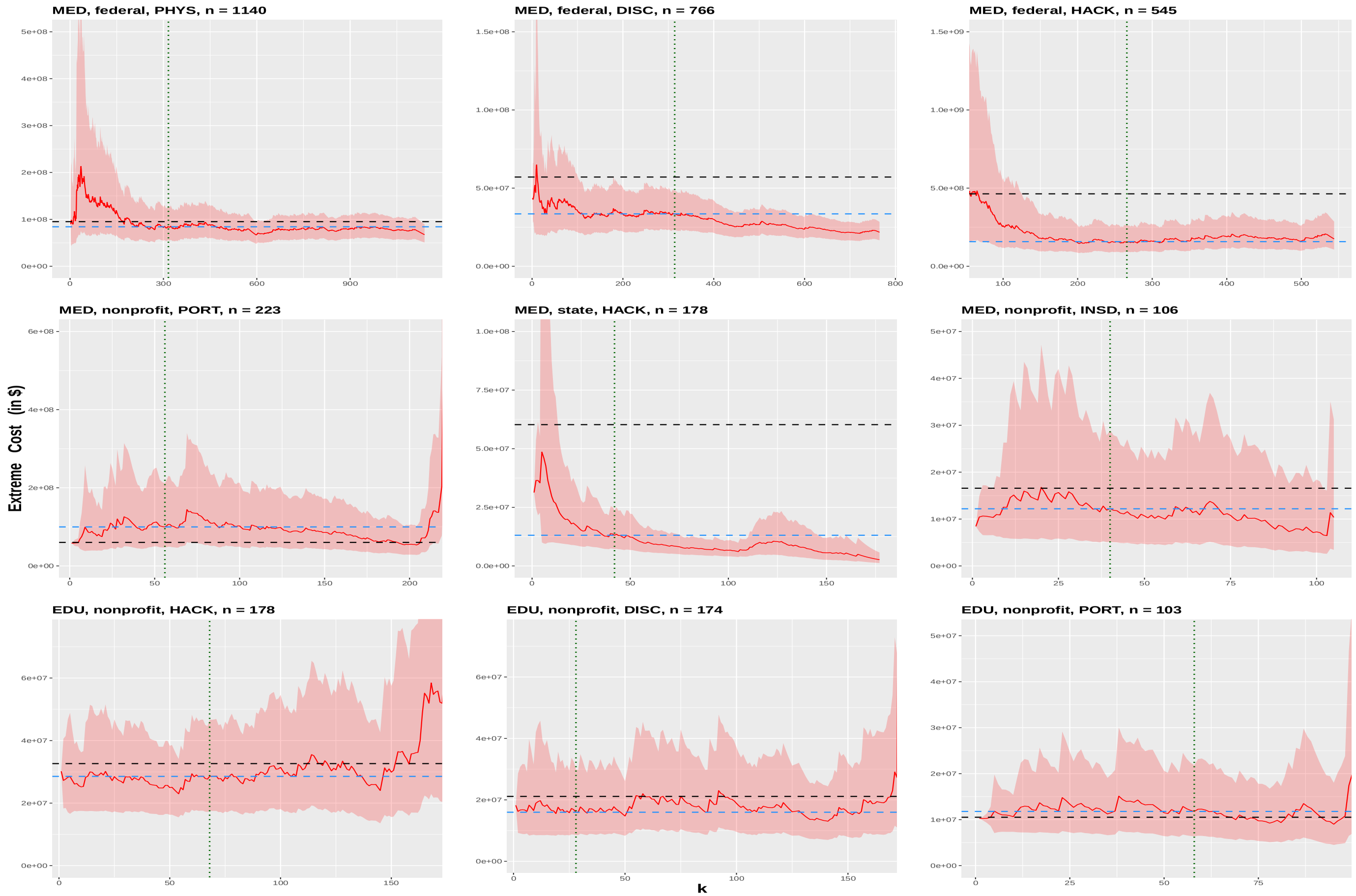

Fig. D30: For each cluster in Table 2, the corresponding plot of the Weissman quantile estimator $\widehat{q}^{\star,\mathrm{BR}}_{1-1/n}$ (calculated by making use of the bias-reduced Hill estimator) versus k (red line), along with its associated asymptotic Gaussian 95% confidence interval (lighter red bands), the value of k selected by the path stability procedure (vertical dotted line) with its corresponding pointwise estimate (horizontal dashed blue line), and the sample maximum (horizontal dashed black line).

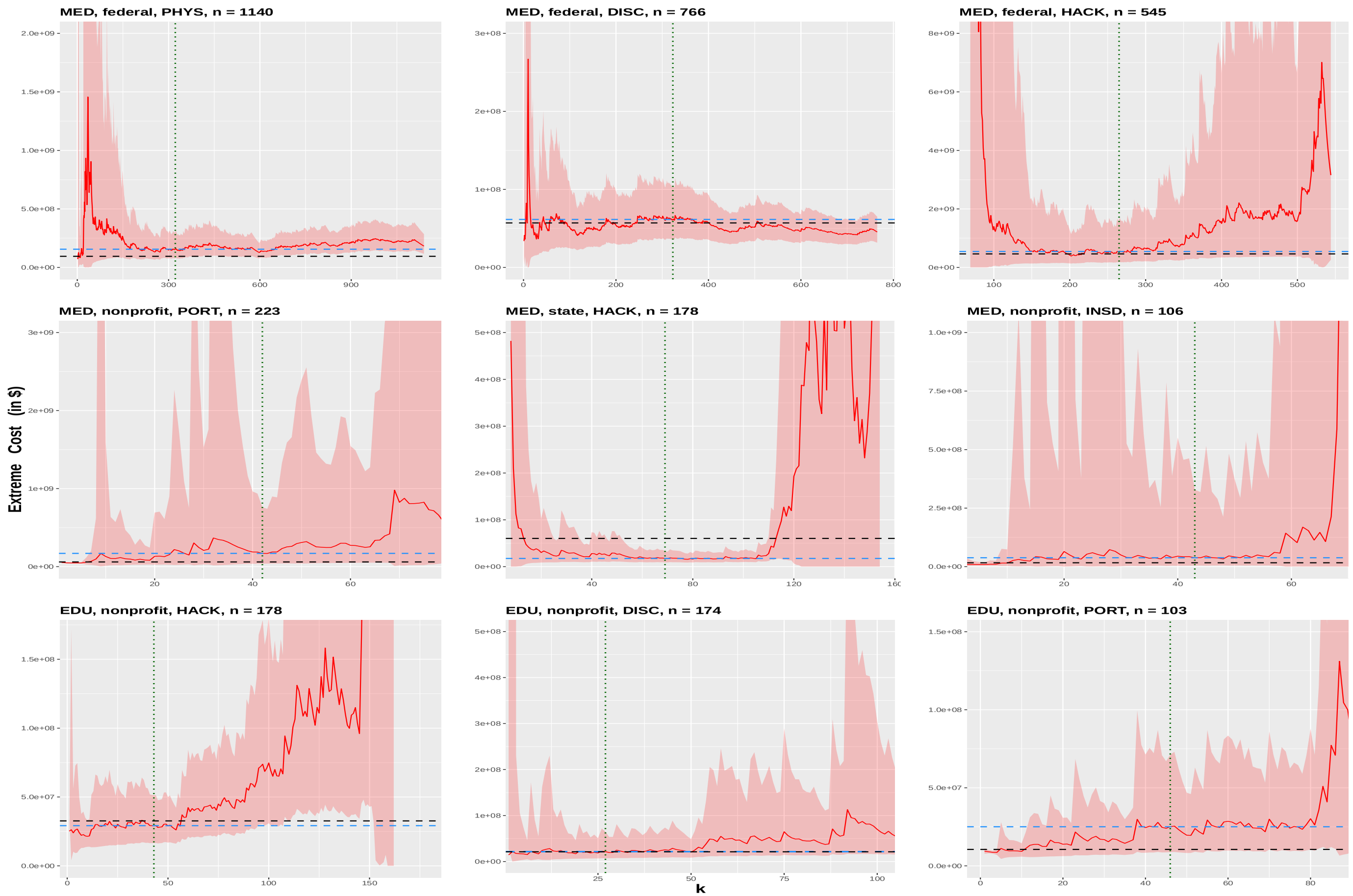

Fig. D31: For each cluster in Table 2, the corresponding plot of the quantile-based bias-reduced expectile estimator $\widetilde{\xi}^{\star,\mathrm{BR}}_{1-1/n}$ versus k (red line), along with the bias- and variance-corrected asymptotic Gaussian 95% confidence interval (lighter red bands), the value of k selected by the path stability procedure (vertical dotted line) with its associated pointwise estimate (horizontal dashed blue line), and the sample maximum (horizontal dashed black line).