On automatic bias reduction for extreme expectile estimation

Stéphane Girard, Gilles Stupfler, Antoine Usseglio-Carleve

On automatic bias reduction for extreme expectile estimation

Stéphane Girard[a], Gilles Stupfler[b] & Antoine Usseglio-Carleve[b]

[a] Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France
[b] Univ Rennes, Ensai, CNRS, CREST - UMR 9194, F-35000 Rennes, France

Abstract

Expectiles induce a law-invariant risk measure that has recently gained popularity in actuarial and financial risk management applications. Unlike quantiles or the quantile-based Expected Shortfall, the expectile risk measure is coherent and elicitable. The estimation of extreme expectiles in the heavy-tailed framework, which is reasonable for extreme financial or actuarial risk management, is not without difficulties; currently available estimators of extreme expectiles are typically biased and hence may show poor finite-sample performance even in fairly large samples. We focus here on the construction of bias-reduced extreme expectile estimators for heavy-tailed distributions. The rationale for our construction hinges on a careful investigation of the asymptotic proportionality relationship between extreme expectiles and their quantile counterparts, as well as of the extrapolation formula motivated by the heavy-tailed context. We accurately quantify and estimate the bias incurred by the use of these relationships when constructing extreme expectile estimators. This motivates the introduction of classes of bias-reduced estimators whose asymptotic properties are rigorously shown, and whose finite-sample properties are assessed on a simulation study and three samples of real data from economics, insurance and finance.

Keywords: Asymmetric least squares, Bias reduction, Expectiles, Extremes, Extrapolation, Heavy tails, Second-order parameter.

1 Introduction

The αth expectile ξ_α of an integrable random variable Y is defined as

$$\xi_\alpha = \arg\min_{\theta\in\mathbb{R}} \mathbb{E}(\eta_\alpha(Y-\theta) - \eta_\alpha(Y)), \tag{1}$$

where $\eta_\alpha(u) = |\alpha - \mathbb{1}\{u \leq 0\}|u^2$ is the so-called expectile check function and $\mathbb{1}\{\cdot\}$ the indicator function. Expectiles are L^2-analogues of quantiles, which are obtained by minimising asymmetrically weighted mean absolute deviations (Koenker and Bassett, 1978):

$$q_\alpha \in \arg\min_{q\in\mathbb{R}} \mathbb{E}(\rho_\alpha(Y-q) - \rho_\alpha(Y)),$$

where $\rho_\alpha(u) = |\alpha - \mathbb{1}\{u \leq 0\}||u|$ is the quantile check function. Expectiles, originally introduced by Newey and Powell (1987) in the context of testing for homoscedasticity and conditional symmetry of the error distribution in linear regression, are always uniquely defined by their convex optimisation problem (unlike quantiles, for which uniqueness is only guaranteed if the underlying distribution function is strictly increasing). They satisfy

$$\alpha = \mathbb{E}[|Y-\xi_\alpha|\mathbb{1}\{Y \leq \xi_\alpha\}]/\mathbb{E}|Y-\xi_\alpha|. \tag{2}$$

In particular, contrary to quantiles, expectiles are determined by tail expectations rather than tail probabilities. Expectiles induce risk measures which have recently gained substantial traction in the risk

management context, for several axiomatic and practical reasons, including the fact that they induce the only risk measure, apart from the simple expectation, defining a law-invariant, coherent (Artzner et al., 1999) and elicitable (Gneiting, 2011) risk measure, see Bellini et al. (2014) and Ziegel (2016). As such, a natural backtesting methodology exists for expectiles. Quantiles are indeed elicitable, but not coherent in general, and are often criticised for missing out on relevant information about distribution tails because their calculation only depends on the frequency of tail events. The Expected Shortfall, meanwhile, takes into account the actual values of the risk variable on the tail event and is a coherent risk measure, but is not elicitable. In financial applications specifically, expectiles are linked through Formula (2) to the notion of gain-loss ratio, which is well-known in the literature on no good deal valuation in incomplete markets and is a popular performance measure in portfolio management (see Bellini and Di Bernardino, 2017, and references therein). Further axiomatic, theoretical and practical justification for the use of expectiles alongside or instead of the quantile and Expected Shortfall can be found in, among others, Ehm et al. (2016) and Bellini and Di Bernardino (2017).

Expectile estimation was first considered in Newey and Powell (1987) in the context of linear regression, and has been developing since then; recent contributions include Sobotka and Kneib (2012) as well as Holzmann and Klar (2016) and Krätschmer and Zähle (2017) for the estimation of central, non-tail expectiles of fixed order α. By contrast, probabilistic aspects of extreme expectiles, with $\alpha \uparrow 1$, were first considered by Bellini et al. (2014) and later Bellini and Di Bernardino (2017). The estimation of extreme expectiles has been considered even more recently in Daouia et al. (2018, 2019, 2020), where it is shown that extreme expectiles can be estimated in several ways. The construction of each of the estimators uses a combination of the heavy-tailed distributional assumption (representing the tail structure of many financial and actuarial data examples fairly well, see *e.g.* p.9 of Embrechts et al. (1997) and p.1 of Resnick (2007)) and a remarkable asymptotic proportionality relationship linking extreme expectiles to their quantile counterparts. An inspection of the finite-sample results of Daouia et al. (2018, 2020) reveals that these estimators suffer from substantial finite-sample bias, even though this is not clear from the asymptotic normality results presented therein. This is of course an issue if expectiles are to be used widely in the management of extreme risk. Partial answers to this bias problem are presented in Girard et al. (2020a) and Girard et al. (2020b) in a regression setup; however, the arguments therein rely on either a simplified setting where one of the leading bias terms can be shown to vanish, or on a crude asymptotic expansion of the bias term where only the part of the bias that is in some sense easiest to correct in the conditional setup is removed. To be more specific, the method of Girard et al. (2020b) is designed to eliminate the source of bias due to the amount of tail heaviness (which has an important influence in the asymptotic proportionality relationship), but cannot handle the bias purely due to the second-order framework. It will therefore perform poorly when this particular source of bias dominates, that is, when the underlying heavy-tailed distribution is far from the standard Pareto distribution on which the extrapolation procedure is based.

The contribution of this paper is to provide a wide class of automatic, data-driven, second-order fully bias-reduced versions of the extreme expectile estimators currently available in the literature. This is done in three steps. First, we briefly recall the construction of extreme expectile estimators at a level $\beta = \beta_n \to 1$ such that $n \to \infty$, where n denotes sample size. This construction is based on the extrapolation of purely empirical expectile estimators at a much lower, intermediate level $\alpha_n \to 1$, with the help of an appropriate tail index estimator, and we highlight how bias may appear from tail index estimation and from the extrapolation procedure itself through a specific bias term. Second, in the Hall-Welsh subclass of heavy-tailed models (Hall and Welsh, 1985) which contains most of the heavy-tailed distributions typically encountered in extreme value analysis, we provide estimators of this bias term that we then use to define versions of extrapolated extreme expectile estimators that are fully corrected for extrapolation bias. Third and last, we discuss the use of bias-reduced estimators of the tail index as a way to complete the elimination of the bias of our extrapolated estimators. We compare the use of an existing bias-reduced version of the Hill estimator constructed in Caeiro et al. (2005) with that of a novel bias-reduced modification of an expectile-based estimator, the biased version having been studied in Girard et al. (2020b) and Padoan and Stupfler (2020). As we shall

see in our simulation study, the expectile-based tail index estimator allows us to gain accuracy mostly in those difficult situations where the so-called second-order parameter is close to 0, that is, when the underlying heavy-tailed distribution is far from the standard Pareto distribution on which the extrapolation procedure is based. In this sense, we make substantial further gains compared to the partial bias-correction procedure discussed in Girard et al. (2020b) that cannot handle this case. The combination of these second and third steps results in fully bias-reduced classes of extrapolated extreme expectile estimators in our heavy-tailed setting. To make these estimators completely automatic, we introduce a selection rule of the Asymptotic Mean Squared Error-optimal value of the tuning parameter α_n representing the upper sample fraction used in our tail index and expectile estimators as well as in the extrapolation bias correction term. This results in estimators whose finite-sample performance is far superior to that of previously considered estimators in the extreme expectile estimation literature, as we shall illustrate in our simulation study.

The paper is organised in the following way. Section 2 gives details on our estimation framework and briefly reviews currently available extreme expectile estimators. Section 3 contains the main contributions of the paper on the bias-reduced estimation of extreme expectiles; all our methods and samples of real data are incorporated into the R package `Expectrem`, currently available at `https://github.com/AntoineUC/Expectrem`, and Section 3.5 explains in detail what this package contains and how to call our methods. Section 4 showcases the performance of our estimators on a simulation study. We finally illustrate the practical applicability of our procedures on real samples of economic, actuarial and financial data in Section 5. More details about the implementation of our methods, mathematical proofs and a complete set of numerical results from our simulation study are relegated to the Supplementary Material document.

2 State of the art on extreme expectile estimation

We start by describing the existing techniques in extreme expectile estimation. Suppose throughout that the available data $(Y_1, \ldots, Y_n)$ is made of independent realisations of the random variable Y with cumulative distribution function F (resp. survival function $\overline{F} = 1 - F$). It is assumed that $\mathbb{E}|Y| < \infty$, so that expectiles of Y of any order exist indeed. Our goal is to estimate an extreme expectile of Y, *i.e.* whose order tends to 1 as $n \to \infty$.

Intermediate level We start by the case of a so-called intermediate level $\alpha_n \to 1$, namely such that $n(1-\alpha_n) \to \infty$ as $n \to \infty$. Intermediate levels tend to infinity slowly enough that expectiles are well within the sample and can thus be estimated by purely empirical methods. It was observed by Jones (1994) that the α_nth expectile is actually the quantile of level α_n associated with the distribution function E defined by

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

Recall that the quantile at level α_n of the distribution function F is defined as $q_{\alpha_n} = \inf\{y \in \mathbb{R} \,|\, F(y) \geq \alpha_n\} = \inf\{y \in \mathbb{R} \,|\, \overline{F}(y) \leq 1-\alpha_n\}$. Intermediate quantiles of F may then be estimated by inverting the empirical conditional survival function induced by the observations:

$$\widehat{q}_{\alpha_n} = \inf\left\{y \in \mathbb{R} \,|\, \widehat{\overline{F}}_n(y) \leq 1-\alpha_n\right\} = Y_{n-\lfloor n(1-\alpha_n)\rfloor,n} \quad \text{with} \quad \widehat{\overline{F}}_n(y) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{Y_i>y\}}.$$

Here $Y_{1,n} \leq Y_{2,n} \leq \cdots \leq Y_{n,n}$ are the order statistics associated with $(Y_1, \ldots, Y_n)$ and $\lfloor \cdot \rfloor$ denotes the floor function. We apply the same principle to the estimation of the intermediate expectile: replacing population averaging with sample averaging in the definition of the distribution function E results in the estimator

$$\widehat{\xi}_{\alpha_n} = \inf\left\{y \in \mathbb{R} \,|\, \widehat{\overline{E}}_n(y) \leq 1-\alpha_n\right\} \quad \text{with} \quad \widehat{\overline{E}}_n(y) = \frac{\sum_{i=1}^{n}(Y_i-y)\mathbb{1}_{\{Y_i>y\}}}{\sum_{i=1}^{n}|Y_i-y|}.$$

This estimator is an unconditional version of the intermediate conditional expectile estimator introduced in Girard et al. (2020b). A straightforward calculation shows that this estimator is in fact also exactly the *Least Asymmetrically Weighted Squares (LAWS) estimator* studied in Daouia et al. (2018), that is, the unique solution of the empirical counterpart of the minimisation problem (1):

$$\widehat{\xi}_{\alpha_n} = \arg\min_{\theta\in\mathbb{R}} \sum_{i=1}^{n} \eta_{\alpha_n}(Y_i - \theta). \tag{3}$$

An alternative estimator can be found, in the class of heavy-tailed distributions we shall focus on hereafter. Recall that the distribution of Y is heavy-tailed if and only if there exists $\gamma > 0$ such that

$$\forall y > 0,\ \lim_{t\to\infty} \frac{\overline{F}(ty)}{\overline{F}(t)} = y^{-1/\gamma} \ \text{ or equivalently } \ \lim_{t\to\infty} \frac{q_{1-(ty)^{-1}}}{q_{1-t^{-1}}} = y^{\gamma}.$$

The tail index γ tunes the tail heaviness of the distribution of Y: if $\gamma > a$ then $\mathbb{E}(Y^{1/a}\mathbb{1}_{\{Y>0\}}) = \infty$ (a precise statement is Exercise 1.16 p.35 in de Haan and Ferreira, 2006). Our minimal working assumption throughout will therefore be that $\gamma < 1$ and $\mathbb{E}(Y_-) < \infty$, where $Y_- = \max(-Y, 0)$. In this case, we have the following asymptotic proportionality relationship between expectile and quantile:

$$\lim_{\alpha\uparrow 1} \frac{\xi_\alpha}{q_\alpha} = (\gamma^{-1} - 1)^{-\gamma}. \tag{4}$$

This was first noted by Bellini et al. (2014). This connection suggests the class of *indirect estimators*

$$\overline{\xi}_{\alpha_n} = (\overline{\gamma}^{-1} - 1)^{-\overline{\gamma}} \widehat{q}_{\alpha_n} = (\overline{\gamma}^{-1} - 1)^{-\overline{\gamma}} Y_{n-\lfloor n(1-\alpha_n)\rfloor,n} \tag{5}$$

where $\overline{\gamma}$ is a consistent estimator of γ.

Extreme level The problem of most relevance in extreme value analysis is to consider the case of a level $\beta_n \to 1$ such that $n(1-\beta_n) \to c < \infty$ as $n \to \infty$. In this situation, purely empirical methods are no longer consistent, and one has to use information about the tail of the data in order to construct an extrapolation procedure. In the context of expectile estimation, this is made possible by the heavy tail assumption and convergence (4): these entail

$$\frac{\xi_{\beta_n}}{\xi_{\alpha_n}} \approx \frac{q_{\beta_n}}{q_{\alpha_n}} \approx \left(\frac{1-\beta_n}{1-\alpha_n}\right)^{-\gamma} \quad \text{as } n \to \infty.$$

We call this approximation the *Weissman approximation*, after the work of Weissman (1978) on extreme quantile estimation. This justifies introducing the class of semiparametric extrapolating estimators

$$\overline{\xi}^{\star}_{\beta_n} = \left(\frac{1-\beta_n}{1-\alpha_n}\right)^{-\overline{\gamma}} \overline{\xi}_{\alpha_n} \tag{6}$$

where $\overline{\xi}_{\alpha_n}$ is any consistent estimator of ξ_{α_n}. One immediately deduces from (6) two subclasses of estimators, replacing $\overline{\xi}_{\alpha_n}$ by the LAWS estimator of ξ_{α_n} or its indirect counterpart; in the latter, the estimator of γ can be chosen different from the estimator featured in the above extrapolation procedure, although we shall not pursue this here for the sake of simplicity. One may also construct a weighted combination of the LAWS-based and indirect estimators, as done in Daouia et al. (2021), although the finite-sample benefit of doing so can be marginal.

The extrapolated LAWS-based and indirect estimators unfortunately suffer from a sizeable amount of finite-sample bias. This is clear from, among others, Figures 3 and 4 in Daouia et al. (2018), where it can be seen that even for the sample size $n = 1{,}000$, and for certain distributions of interest in extreme value modelling, these estimators have a relative bias of the order of 50%. This means that the estimator is on average 50% larger than the target extreme expectile. The contributions of this paper, which we gather in the next section, are a precise quantification of this bias using a standard second-order refinement of the heavy tail condition, and the introduction of automatic bias reduction procedures whose finite-sample performance will be examined in detail in Sections 4 and 5.

3 Automatic bias reduction methodology for extreme expectile estimation

3.1 Rationale for our bias correction methods

The construction of the class of extrapolated estimators in (6) relies on the successive use of Equation (4), in order to approximate a ratio of high expectiles by a ratio of high quantiles at corresponding levels, and an approximation of this ratio of high quantiles that is warranted by the heavy tail assumption. The magnitude of the bias of the extrapolated estimators will therefore be crucially driven by the rates of convergence and the error terms in these two approximations.

To simplify the exposition, we assume from now on that $k_n = n(1-\alpha_n)$ is a sequence of positive integers, and we rewrite the assumptions $\alpha_n \to 1$ and $n(1-\alpha_n) \to \infty$ as $k_n \to \infty$ and $k_n/n \to 0$. This choice is motivated by the fact that in quantile estimation, this quantity k_n denotes the effective sample size, *i.e.* the number of top order statistics eventually used for the estimation. Adopting this convention will make it easier to state and compare our results with existing results in the extreme value analysis of heavy tails. This is not a restriction in practice: our estimators of properly extreme expectiles, having order $\beta_n \to 1$ such that $n(1-\beta_n) \to c < \infty$, are built on intermediate expectile estimators whose level α_n we are free to choose, and those levels such that $k_n = n(1-\alpha_n)$ is an integer constitute a sufficient set of levels to work with.

A classical device in extreme value analysis for bias quantification is the following second-order regular variation condition that refines our initial heavy tail assumption.

$\mathcal{C}_2(\gamma, \rho, A)$ The survival function $\overline{F}$ is second-order regularly varying with index $-1/\gamma < 0$, second-order parameter $\rho \leq 0$ and a measurable auxiliary function A having constant sign and converging to 0 at infinity, *i.e.*

$$\forall y > 0,\ \lim_{t\to\infty} \frac{1}{A(1/\overline{F}(t))} \left(\frac{\overline{F}(ty)}{\overline{F}(t)} - y^{-1/\gamma} \right) = y^{-1/\gamma} \frac{y^{\rho/\gamma} - 1}{\gamma\rho}.$$

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

An equivalent condition on the tail quantile function $t \mapsto q_{1-t^{-1}}$ is that

$$\forall y > 0,\ \lim_{t\to\infty} \frac{1}{A(t)} \left[\frac{q_{1-(ty)^{-1}}}{q_{1-t^{-1}}} - y^{\gamma} \right] = y^{\gamma} \frac{y^{\rho} - 1}{\rho}.$$

See de Haan and Ferreira (2006, Theorem 2.3.9). Numerous examples of commonly used distributions that satisfy this assumption can be found in Beirlant et al. (2004).

The fundamental argument behind our methodology is that the direct and indirect extrapolated estimators, *i.e.*

$$\widehat{\xi}^{\star}_{\beta_n} = \left(\frac{n(1-\beta_n)}{k_n} \right)^{-\overline{\gamma}} \widehat{\xi}_{1-k_n/n} \quad \text{and} \quad \widetilde{\xi}^{\star}_{\beta_n} = \left(\frac{n(1-\beta_n)}{k_n} \right)^{-\overline{\gamma}} (\overline{\gamma}^{-1} - 1)^{-\overline{\gamma}} Y_{n-k_n,n}$$

where $\overline{\gamma}$ is a $\sqrt{k_n}-$consistent estimator of γ, satisfy

$$\log\left(\frac{\widehat{\xi}^{\star}_{\beta_n}}{\xi_{\beta_n}} \right) = (\overline{\gamma} - \gamma) \log\left(\frac{k_n}{n(1-\beta_n)} \right) + \log\left(\frac{\widehat{\xi}_{1-k_n/n}}{\xi_{1-k_n/n}} \right) - \log\left(\left[\frac{n(1-\beta_n)}{k_n} \right]^{\gamma} \frac{\xi_{\beta_n}}{\xi_{1-k_n/n}} \right) \tag{7}$$

$$\text{and } \log\left(\frac{\widetilde{\xi}^{\star}_{\beta_n}}{\xi_{\beta_n}} \right) = (\overline{\gamma} - \gamma) \log\left(\frac{k_n}{n(1-\beta_n)} \right) + \log\left(\frac{(\overline{\gamma}^{-1} - 1)^{-\overline{\gamma}}}{(\gamma^{-1} - 1)^{-\gamma}} \right) + \log\left(\frac{Y_{n-k_n,n}}{q_{1-k_n/n}} \right)$$
$$- \log\left(\left[\frac{n(1-\beta_n)}{k_n} \right]^{\gamma} (\gamma^{-1} - 1)^{\gamma} \frac{\xi_{\beta_n}}{q_{1-k_n/n}} \right). \tag{8}$$

Under condition $\mathcal{C}_2(\gamma,\rho,A)$ and standard technical assumptions on k_n and β_n, $\widehat{\xi}_{1-k_n/n}$ and $Y_{n-k_n,n}$ have the same rate of convergence $1/\sqrt{k_n}$, which is also the rate of convergence of the final (pure bias) nonrandom term, see *e.g.* Theorem 5 in Daouia et al. (2020) and its proof. Since $\log(k_n/[n(1-\beta_n)]) \to \infty$, the first term dominates, leading to a common asymptotic distribution for $\widehat{\xi}^{\star}_{\beta_n}$ and $\widetilde{\xi}^{\star}_{\beta_n}$: if $\sqrt{k_n}(\overline{\gamma}-\gamma) \xrightarrow{\mathrm{d}} \Gamma$, then

$$\frac{\sqrt{k_n}}{\log(k_n/[n(1-\beta_n)])}\log\left(\frac{\widehat{\xi}^{\star}_{\beta_n}}{\xi_{\beta_n}}\right) \xrightarrow{\mathrm{d}} \Gamma \quad \text{and} \quad \frac{\sqrt{k_n}}{\log(k_n/[n(1-\beta_n)])}\log\left(\frac{\widetilde{\xi}^{\star}_{\beta_n}}{\xi_{\beta_n}}\right) \xrightarrow{\mathrm{d}} \Gamma. \tag{9}$$

A naive bias-correction strategy would thus focus on the bias incurred in the estimation of γ, identifiable through the asymptotic distribution Γ. However, Equations (7) and (8) reveal another source of bias, which is the use of the Weissman approximation itself to control the final terms in Equations (7) and (8) (note that the second term in Equation (7) and the third term in Equation (8) are asymptotically unbiased, see Theorem 2 in Daouia et al. (2018) and Theorem 2.4.8 p.52 in de Haan and Ferreira (2006)). Our contribution hereafter is to design fully bias-reduced extreme expectile estimators by eliminating both of these sources of bias.

3.2 Construction of bias-reduced extreme expectile estimators

We construct bias-reduced versions of the extrapolated estimators $\widehat{\xi}^{\star}_{\beta_n}$ and $\widetilde{\xi}^{\star}_{\beta_n}$ in two steps. Our methods will feature estimators of the second-order parameter ρ, but also estimators of the auxiliary function A. Estimating this function without any further assumption can be a difficult task; however, for most of the distributions satisfying condition $\mathcal{C}_2(\gamma,\rho,A)$ used for modelling purposes, the function A takes the form $A(t)=b\gamma t^{\rho}$, for a certain nonzero constant b and $\rho<0$. We assume in what follows that the function A is indeed of this form, which amounts to assuming that the underlying distribution belongs to the Hall-Welsh class in the sense of Gomes and Pestana (2007). We give a list of examples of classical heavy-tailed distributions in Table 1, containing among others the distributions we shall work with in our simulation study, with their respective values of γ, ρ and b.

The function A can then be estimated using consistent estimators $\overline{\gamma}$, $\overline{b}$ and $\overline{\rho}$ of γ, b and ρ, respectively. We assume in Section 3.2.1 that such estimators are given; we shall explain in detail in Section 3.2.2 which estimators $\overline{\gamma}$ we consider, while the estimators $\overline{b}$ and $\overline{\rho}$ are calculated directly from the R package `mop` (see Appendix A for a brief summary of how these estimators are constructed). We start by dealing with the bias due to the extrapolation procedure itself, contained in the final terms of Equations (7) and (8).

3.2.1 Bias due to the extrapolation procedure

We start by the nonrandom bias term in Equation (7), and we write

$$\left[\frac{n(1-\beta_n)}{k_n}\right]^{\gamma}\frac{\xi_{\beta_n}}{\xi_{1-k_n/n}} = \underbrace{\left[\frac{n(1-\beta_n)}{k_n}\right]^{\gamma}\frac{q_{\beta_n}}{q_{1-k_n/n}}}_{1+B_{1,n}} \times \underbrace{\left(\gamma^{-1}-1\right)^{-\gamma}\frac{q_{1-k_n/n}}{\xi_{1-k_n/n}}}_{1+B_{2,n}} \times \underbrace{\left(\gamma^{-1}-1\right)^{\gamma}\frac{\xi_{\beta_n}}{q_{\beta_n}}}_{1+B_{3,n}}. \tag{10}$$

By Theorem 2.3.9 in de Haan and Ferreira (2006), the bias term $B_{1,n}$ can be written as

$$B_{1,n} = \frac{[n(1-\beta_n)/k_n]^{-\rho}-1}{\rho}A(n/k_n)(1+\mathrm{o}(1)).$$

We now focus on the other two bias terms $B_{2,n}$ and $B_{3,n}$ linking an expectile to its quantile counterpart, at intermediate and extreme levels. It follows from the proof of Proposition 1 in Daouia et al. (2018)

that

$$\frac{\overline{F}(\xi_\alpha)}{1-\alpha} = \left(\gamma^{-1}-1\right)(1+r(\alpha)) \tag{11}$$

$$\text{with } 1+r(\alpha) = \left(1-\frac{\mathbb{E}[Y]}{\xi_\alpha}\right)\frac{1}{2\alpha-1}\left(1+A\left(\frac{1}{\overline{F}(\xi_\alpha)}\right)\frac{1}{\gamma(1-\gamma-\rho)}(1+\mathrm{o}(1))\right)^{-1} \quad \text{as } \alpha\uparrow 1.$$

Using Lemma 1 in Daouia et al. (2020) together with the heavy tail assumption then entails

$$\frac{\xi_\alpha}{q_\alpha} = \left(\gamma^{-1}-1\right)^{-\gamma}(1+r(\alpha))^{-\gamma}\left(1+\frac{\left(\gamma^{-1}-1\right)^{-\rho}(1+r(\alpha))^{-\rho}-1}{\rho}A((1-\alpha)^{-1})(1+\mathrm{o}(1))\right) \tag{12}$$

as $\alpha\to 1$. With $\alpha = 1-k_n/n$ and $\alpha=\beta_n$, this yields

$$B_{2,n} = (1+r(1-k_n/n))^{\gamma}\left(1+\frac{\left(\gamma^{-1}-1\right)^{-\rho}(1+r(1-k_n/n))^{-\rho}-1}{\rho}A(n/k_n)(1+\mathrm{o}(1))\right)^{-1}-1$$

$$\text{and } B_{3,n} = (1+r(\beta_n))^{-\gamma}\left(1+\frac{\left(\gamma^{-1}-1\right)^{-\rho}(1+r(\beta_n))^{-\rho}-1}{\rho}A((1-\beta_n)^{-1})(1+\mathrm{o}(1))\right)-1.$$

Each of these bias terms can be estimated. Recall our assumption that $A(t)=b\gamma t^\rho$, and use consistent estimators $\overline{\gamma}$ of γ, $\overline{\rho}$ of ρ and $\overline{b}$ of b to estimate $B_{1,n}$ by

$$\overline{B}_{1,n} = \frac{[n(1-\beta_n)/k_n]^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(n/k_n)^{\overline{\rho}}.$$

Let further $\overline{Y}_n$ denote the sample mean of $Y_1,\ldots,Y_n$, $\overline{\xi}_{1-k_n/n}$ be either the LAWS or indirect intermediate expectile estimator, and $\overline{\xi}^\star_{\beta_n}$ be the related extrapolated estimator (in our current implementation we use the LAWS estimator for $\overline{\xi}_{1-k_n/n}$ and its extrapolated version for $\overline{\xi}^\star_{\beta_n}$). The remainder terms $r(1-k_n/n)$ and $r(\beta_n)$ are estimated by

$$\overline{r}(1-k_n/n) = \left(1-\frac{\overline{Y}_n}{\overline{\xi}_{1-k_n/n}}\right)\frac{1}{1-2k_n/n}\left(1+\frac{\overline{b}[\widehat{\overline{F}}_n(\overline{\xi}_{1-k_n/n})]^{-\overline{\rho}}}{1-\overline{\gamma}-\overline{\rho}}\right)^{-1}-1$$

$$\text{and } \overline{r}^\star(\beta_n) = \left(1-\frac{\overline{Y}_n}{\overline{\xi}^\star_{\beta_n}}\right)\frac{1}{2\beta_n-1}\left(1+\frac{\overline{b}\left(\overline{\gamma}^{-1}-1\right)^{-\overline{\rho}}}{1-\overline{\gamma}-\overline{\rho}}(1-\beta_n)^{-\overline{\rho}}\right)^{-1}-1.$$

This yields estimators of $B_{2,n}$ and $B_{3,n}$ as

$$\overline{B}_{2,n} = (1+\overline{r}(1-k_n/n))^{\overline{\gamma}}\left(1+\frac{\left(\overline{\gamma}^{-1}-1\right)^{-\overline{\rho}}(1+\overline{r}(1-k_n/n))^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(n/k_n)^{\overline{\rho}}\right)^{-1}-1$$

$$\text{and } \overline{B}_{3,n} = (1+\overline{r}^\star(\beta_n))^{-\overline{\gamma}}\left(1+\frac{\left(\overline{\gamma}^{-1}-1\right)^{-\overline{\rho}}(1+\overline{r}^\star(\beta_n))^{-\overline{\rho}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(1-\beta_n)^{-\overline{\rho}}\right)-1.$$

We deduce from (7) and (10) that a version of the direct extrapolated estimator $\widehat{\xi}^\star_{\beta_n}$, corrected for the bias exclusively due to the heavy-tailed extrapolation, is

$$\begin{aligned}\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n} &= \widehat{\xi}^\star_{\beta_n}(1+\overline{B}_{1,n})(1+\overline{B}_{2,n})(1+\overline{B}_{3,n})\\ &= \left[\left(\frac{n(1-\beta_n)}{k_n}\right)^{-\overline{\gamma}}\widehat{\xi}_{1-k_n/n}\right](1+\overline{B}_{1,n})(1+\overline{B}_{2,n})(1+\overline{B}_{3,n}).\end{aligned} \tag{13}$$

A correction for the extrapolation bias in the indirect estimator is simpler. Indeed,

$$\left[\frac{n(1-\beta_n)}{k_n}\right]^{\gamma}\left(\gamma^{-1}-1\right)^{\gamma}\frac{\xi_{\beta_n}}{q_{1-k_n/n}}=\underbrace{\left[\frac{n(1-\beta_n)}{k_n}\right]^{\gamma}\frac{q_{\beta_n}}{q_{1-k_n/n}}}_{1+B_{1,n}}\times\underbrace{\left(\gamma^{-1}-1\right)^{\gamma}\frac{\xi_{\beta_n}}{q_{\beta_n}}}_{1+B_{3,n}}. \tag{14}$$

From (14), a version of the indirect extrapolated estimator $\widetilde{\xi}^{\star}_{\beta_n}$, corrected for the bias exclusively due to the heavy-tailed extrapolation, is then

$$\begin{aligned}\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}&=\widetilde{\xi}^{\star}_{\beta_n}(1+\overline{B}_{1,n})(1+\overline{B}_{3,n})\\&=\left[\left(\frac{n(1-\beta_n)}{k_n}\right)^{-\overline{\gamma}}(\overline{\gamma}^{-1}-1)^{-\overline{\gamma}}Y_{n-k_n,n}\right](1+\overline{B}_{1,n})(1+\overline{B}_{3,n}).\end{aligned} \tag{15}$$

The methodology introduced here differs from the earlier bias reduction techniques introduced in Girard et al. (2020a) and Girard et al. (2020b) in a regression setup. In Girard et al. (2020a), the bias term proportional to $1/\xi_{1-k_n/n}$ is not corrected because the theory therein focuses on a random variable with expectation 0; in Girard et al. (2020b), the bias term proportional to $A(n/k_n)$ is not corrected because it is very difficult to correct accurately this source of bias in the conditional, nonparametric setup on which that paper focuses. In addition, the correction terms in the aforementioned papers rely on linearising the bias terms, whereas we keep the structure of the bias as intact as possible. This makes our correction term $B_{2,n}$ more accurate indeed than in these earlier attempts. The inclusion of term $B_{3,n}$ is also new; while it could be expected that this term only has a small influence because it relies on quantities calculated at a higher asymptotic order, it is our experience that its inclusion substantially improves finite-sample performance.

We now concentrate on reducing the bias due to the estimation of γ. Combined with the general corrections in (13) and (15), this will result in a fully bias-corrected extrapolated estimator.

3.2.2 Bias reduction for tail index estimation: an expectile-based method

Numerous tail index estimators have been introduced and studied in the literature; a review of some of the most important estimators is given in de Haan and Ferreira (2006, Chapter 3). There are various techniques for the reduction of bias of such estimators, an excellent summary being given in the Introduction of Cai et al. (2013). Here our contribution is to propose a bias-reduced version of a purely expectile-based tail index estimator, our procedure being partly inspired by a method developed in, among others, Caeiro et al. (2005). To make the construction of this estimator easier, we start by briefly recalling how the technique of Caeiro et al. (2005) works. Consider the classical Hill estimator (Hill, 1975):

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

It is known that under $\mathcal{C}_2(\gamma,\rho,A)$, if moreover $\sqrt{k_n}A(n/k_n)\to\lambda\in\mathbb{R}$, then (see Theorem 3.2.5 in de Haan and Ferreira, 2006)

$$\sqrt{k_n}(\widehat{\gamma}^{\mathrm{H}}_{k_n}-\gamma)\xrightarrow{\mathrm{d}}\mathcal{N}\left(\frac{\lambda}{1-\rho},\gamma^2\right).$$

In finite samples $\lambda\approx\sqrt{k_n}A(n/k_n)=\sqrt{k_n}b\gamma(n/k_n)^{\rho}$, meaning that the pseudo-estimator

$$\widehat{\gamma}^{\mathrm{H}}_{k_n}\left(1-\frac{b}{1-\rho}\left(\frac{n}{k_n}\right)^{\rho}\right)$$

should be asymptotically unbiased with the same variance as the Hill estimator. Caeiro et al. (2005) then plug in consistent estimators $\overline{b}$ of b and $\overline{\rho}$ of ρ and arrive at the bias-reduced Hill estimator

$$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}} = \widehat{\gamma}_{k_n}^{\mathrm{H}} \left(1 - \frac{\overline{b}}{1-\overline{\rho}} \left(\frac{n}{k_n}\right)^{\overline{\rho}}\right).$$

Theorem 3.1 of Caeiro et al. (2005) shows that $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ is indeed $\sqrt{k_n}-$asymptotically Gaussian with expectation zero and variance γ^2. The construction of this estimator essentially hinges on eliminating the bias by multiplying the original estimator by a quantity cancelling this bias.

We adapt here this construction to propose a bias-reduction procedure for the estimator

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

The rationale behind this estimator, studied in different contexts by Girard et al. (2020b) and Padoan and Stupfler (2020), is that, from (11),

$$\gamma = \left(1 + \frac{\overline{F}(\xi_\alpha)}{1-\alpha}\frac{1}{1+r(\alpha)}\right)^{-1} \approx \left(1 + \frac{\overline{F}(\xi_\alpha)}{1-\alpha}\right)^{-1} \quad \text{as } \alpha \uparrow 1.$$

To find a bias-reduced version of this *asymptotic proportionality tail index estimator* $\widehat{\gamma}_{k_n}^{\mathrm{E}}$, we follow the above idea and use the sample counterpart $\overline{r}(1-k_n/n)$ of $r(1-k_n/n)$ defined in Section 3.2.1: this yields a bias-reduced asymptotic proportionality tail index estimator as

$$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}} = \left(1 + \frac{n\widehat{\overline{F}}_n(\widehat{\xi}_{1-k_n/n})}{k_n}\frac{1}{1+\overline{r}(1-k_n/n)}\right)^{-1}.$$

In our current implementation we take $\overline{\xi}_{1-k_n/n} = \widehat{\xi}_{1-k_n/n}$ and $\overline{\gamma} = \widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ inside $\overline{r}(1-k_n/n)$, specifically for the calculation of this bias-reduced version.

Our first main theoretical result gives the asymptotic normality and unbiasedness of $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$.

Theorem 1. *Suppose that $\mathbb{E}|Y_-|^{2+\delta} < \infty$ for some $\delta > 0$. Assume further that $\mathcal{C}_2(\gamma, \rho, A)$ holds with $0 < \gamma < 1/2$, $\rho < 0$ and $A(t) = b\gamma t^\rho$, and let k_n be a sequence such that $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. If $\sqrt{k_n}A(n/k_n) \to \lambda_1 \in \mathbb{R}$, $\sqrt{k_n}/q(1-k_n/n) \to \lambda_2 \in \mathbb{R}$, and $\overline{\gamma}$, $\overline{\rho}$ and $\overline{b}$ are consistent estimators of γ, ρ and b such that $(\overline{\rho}-\rho)\log(n) = o_{\mathbb{P}}(1)$, then*

$$\sqrt{k_n}(\widehat{\gamma}_{k_n}^{\mathrm{E,RB}} - \gamma) \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \frac{\gamma^3(1-\gamma)}{1-2\gamma}\right).$$

We provide a comparison of the two tail index estimators in terms of variance in Figure 1. The asymptotic variance of $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ is substantially smaller than that of $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ when γ is less than 0.35. The variance of $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ explodes, however, as $\gamma \uparrow 1/2$, which is to be expected since the estimators $\widehat{\gamma}_{k_n}^{\mathrm{E}}$ and $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ are based on the intermediate LAWS estimator $\widehat{\xi}_{1-k_n/n}$, itself known to be asymptotically normal only when $\gamma < 1/2$ (see Daouia et al., 2018). In our implementation (and in particular in the calculation of the $\overline{B}_{j,n}$ in Section 3.2.1), one can choose either $\overline{\gamma} = \widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ or $\overline{\gamma} = \widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$.

3.3 Fully bias-reduced estimators and their asymptotic properties

Our final, fully bias-reduced estimators are members of the following two classes. The first class, which extrapolates the intermediate LAWS estimator, is defined as

$$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}} = \left[\left(\frac{n(1-\beta_n)}{k_n}\right)^{-\overline{\gamma}} \widehat{\xi}_{1-k_n/n}\right](1+\overline{B}_{1,n})(1+\overline{B}_{2,n})(1+\overline{B}_{3,n})$$

where $\overline{\gamma}$ is any bias-reduced estimator of the tail index γ, and $\overline{B}_{1,n}$, $\overline{B}_{2,n}$ and $\overline{B}_{3,n}$ are defined in Section 3.2.1. The second class, based on extrapolating the indirect estimator, is

$$\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n} = \left[\left(\frac{n(1-\beta_n)}{k_n}\right)^{-\overline{\gamma}} (\overline{\gamma}^{-1}-1)^{-\overline{\gamma}} Y_{n-k_n,n}\right](1+\overline{B}_{1,n})(1+\overline{B}_{3,n})$$

where again $\overline{\gamma}$ is any bias-reduced estimator of the tail index γ. It should be noted that this second class of estimators has a nice interpretation in terms of a bias-reduced version for the Weissman extreme quantile estimator (Weissman, 1978; Gomes and Pestana, 2007): this estimator is, with our notation,

$$\widetilde{q}^{\star,\mathrm{RB}}_{\beta_n} = \left[\left(\frac{n(1-\beta_n)}{k_n}\right)^{-\overline{\gamma}} Y_{n-k_n,n}\right](1+\overline{B}_{1,n}).$$

It follows that an equivalent expression for $\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}$ is

$$\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n} = \left[(\overline{\gamma}^{-1}-1)^{-\overline{\gamma}} \widetilde{q}^{\star,\mathrm{RB}}_{\beta_n}\right](1+\overline{B}_{3,n}).$$

In other words, this estimator is obtained by using the asymptotic proportionality relationship (4) at level $\alpha = \beta_n$, plugging in bias-reduced estimators of the tail index and extreme quantile involved, and finally correcting directly for the bias incurred using this proportionality relationship at the level β_n only.

We briefly explore the asymptotic properties of these bias-reduced estimators $\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}$ and $\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}$. It was highlighted in Section 3.1 (see (9)) that extrapolated expectile estimators have a limiting distribution controlled by their tail index estimator. This is also true for their bias-reduced versions, as the following result shows.

Theorem 2. *Assume that $\mathcal{C}_2(\gamma,\rho,A)$ holds with $\rho<0$ and $A(t)=b\gamma t^{\rho}$, and let k_n, β_n be two sequences such that $k_n\to\infty$, $k_n/n\to 0$, $n(1-\beta_n)/k_n\to 0$ and $k_n^{-1/2}\log(k_n/[n(1-\beta_n)])\to 0$ as $n\to\infty$. Assume further that $\sqrt{k_n}A(n/k_n)\to\lambda_1\in\mathbb{R}$, $\sqrt{k_n}/q(1-k_n/n)\to\lambda_2\in\mathbb{R}$, $\sqrt{k_n}(\overline{\gamma}-\gamma)\xrightarrow{\mathrm{d}}\Gamma$ where Γ is a nondegenerate distribution, and $\overline{\rho}$ and $\overline{b}$ are consistent estimators of ρ and b such that $(\overline{\rho}-\rho)\log(n)=\mathrm{o}_{\mathbb{P}}(1)$.*

(i) If $\mathbb{E}|Y_-|<\infty$ and $0<\gamma<1$, then $\dfrac{\sqrt{k_n}}{\log(k_n/[n(1-\beta_n)])}\left(\dfrac{\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}}{\xi_{\beta_n}}-1\right)\xrightarrow{\mathrm{d}}\Gamma.$

(ii) If moreover $\mathbb{E}|Y_-|^2<\infty$ and $0<\gamma<1/2$, then $\dfrac{\sqrt{k_n}}{\log(k_n/[n(1-\beta_n)])}\left(\dfrac{\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}}{\xi_{\beta_n}}-1\right)\xrightarrow{\mathrm{d}}\Gamma.$

It follows from Theorem 2 that the bias-reduced extrapolated expectile estimators have similar asymptotic properties as their standard counterparts. We shall illustrate in Section 4, however, that the bias-reduced versions generally have much better finite-sample properties. Before that, we explain how to select the important tuning parameter k_n appearing in the estimation of the tail index, intermediate expectile level, and bias correction terms.

3.4 Choice of the intermediate level k_n

The choice of the sequence k_n is a crucial point: a low k_n translates into a large variance, and a high k_n translates into a large bias. Choosing k_n therefore leads to solving a trade-off between the bias and variance of the tail index estimator to be used. In order to find the right balance, de Haan and Ferreira (2006) proposed a choice of k_n for the Hill estimator as a minimiser of an estimate of the Asymptotic Mean Squared Error of this estimator. We propose to develop our own selection rule for k_n, to be

used in conjunction with the expectile-based bias-reduced asymptotic proportionality estimator $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$, based on the ordinary asymptotic proportionality estimator $\widehat{\gamma}^{\mathrm{E}}_{k_n}$ also presented in Section 3.2.2. For this estimator, it holds that

$$\sqrt{k_n}\left(\widehat{\gamma}^{\mathrm{E}}_{k_n} - \gamma - \frac{\gamma(\gamma^{-1}-1)^{1-\rho}}{1-\gamma-\rho}A(n/k_n) - \frac{\gamma^2(\gamma^{-1}-1)^{\gamma+1}\mathbb{E}[Y]}{q(1-k_n/n)}\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \frac{\gamma^3(1-\gamma)}{1-2\gamma}\right).$$

See Proposition 1 in Appendix B. Consequently, there are two sources of bias in the expectile-based estimator $\widehat{\gamma}^{\mathrm{E}}_{k_n}$: one proportional to $A(n/k_n)$, having order $(n/k_n)^{\rho}$, and another proportional to $1/q(1-k_n/n)$, having order $(n/k_n)^{-\gamma}$. The leading term of bias will thus depend on ρ and γ. The second source of bias, proportional to $1/q(1-k_n/n)$, can very accurately be eliminated; indeed, its expression only features γ, $\mathbb{E}[Y]$ and $q(1-k_n/n)$, for which we have good estimators that converge to the rate $\sqrt{k_n}$ or more. By contrast, the first source of bias features second-order parameters from the distribution of Y, whose estimators converge slowly (see *e.g.* p.298 in Gomes et al. (2009) and p.2638 in Goegebeur et al. (2010)), and thus is more difficult to remove. In practice, this means that the trade-off to be solved when using the expectile-based asymptotic proportionality tail index estimator will essentially be between the bias due to the second-order quantity $A(n/k_n)$ and the variance of the estimator. This gives us the idea of minimising the *Partial Asymptotic Mean Squared Error*

$$\begin{aligned}\mathrm{PAMSE}(k_n) &= \left[\frac{\gamma(\gamma^{-1}-1)^{1-\rho}}{1-\gamma-\rho}A(n/k_n)\right]^2 + \frac{\gamma^3(1-\gamma)}{1-2\gamma}\times\frac{1}{k_n}\\ &\propto \frac{b^2(\gamma^{-1}-1)^{1-2\rho}}{(1-\gamma-\rho)^2}\times\left(\frac{n}{k_n}\right)^{2\rho} + \frac{1}{1-2\gamma}\times\frac{1}{k_n}.\end{aligned}$$

It is readily seen that this leads to the optimal value

$$k^{\mathrm{E}}_n = \left(\frac{(\gamma^{-1}-1)^{2\rho-1}(1-\gamma-\rho)^2}{-2\rho b^2(1-2\gamma)}\right)^{1/(1-2\rho)} n^{-2\rho/(1-2\rho)}.$$

This optimal value depends on the unknown γ, ρ and b; in practice we use the estimated value

$$\widehat{k}^{\mathrm{E}}_n = \min\left(\left\lfloor\left(\frac{(\overline{\gamma}^{-1}-1)^{2\overline{\rho}-1}(1-\overline{\gamma}-\overline{\rho})^2}{-2\overline{\rho}\overline{b}^2(1-2\overline{\gamma})}\right)^{1/(1-2\overline{\rho})} n^{-2\overline{\rho}/(1-2\overline{\rho})}\right\rfloor, \left\lfloor\frac{n}{2}\right\rfloor - 1\right). \tag{16}$$

In Equation (16), the estimators $\overline{\rho}$ and $\overline{b}$ are the same as in Section 3.2, and $\overline{\gamma}$ is the bias-reduced Hill estimator $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ with the corresponding estimated AMSE-optimal choice of k_n, that is

$$\widehat{k}^{\mathrm{H}}_n = \left\lfloor\left(\frac{(1-\overline{\rho})^2}{-2\overline{\rho}\overline{b}^2}\right)^{1/(1-2\overline{\rho})} n^{-2\overline{\rho}/(1-2\overline{\rho})}\right\rfloor.$$

The fact that we force our selected $\widehat{k}^{\mathrm{E}}_n$ to be less than $\lfloor n/2\rfloor$ is due to the presence of the multiplicative term $(1-2k_n/n)^{-1}$ in our bias reduction methodology (featuring in $\overline{r}(1-k_n/n)$). Since $n^{-2\rho/(1-2\rho)} = \mathrm{o}(n)$ for any $\rho < 0$, this restriction disappears with arbitrarily high probability as $n \to \infty$. We recommend this choice $\widehat{k}^{\mathrm{E}}_n$ when the expectile-based asymptotic proportionality estimator $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ is used.

3.5 The `Expectrem` package

We have implemented our methods in an R package called `Expectrem`, freely downloadable at `https://github.com/AntoineUC/Expectrem`. The content of this package is the following:

- Basic functions for the estimation: `Fbarhat` returns the empirical estimator of the survival function, and `expect` provides the empirical LAWS expectile estimator at a given level. Basic population expectile calculations: `enorm`, `et`, `elog`, `epareto`, `egpd` and `eburr` respectively return the expectiles of the normal, Student, logistic, Pareto, Generalised Pareto (GP) and Burr distributions.
- Tail index estimation: The estimator $\widehat{\gamma}_{k_n}^{\mathrm{E}}$ is implemented in the function `tindexp` with argument `br=FALSE` (default). If `br=TRUE`, the bias-reduced estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ is returned. The other optional argument is the intermediate level `k`, set at $k = \widehat{k}_n^{\mathrm{E}}$ by default.
- Extreme expectile estimation: The direct and indirect extreme expectile estimators $\widehat{\xi}_{\beta_n}^{\star}$ and $\widetilde{\xi}_{\beta_n}^{\star}$ are computed in the function `extExpect` with arguments `method="direct"` and `method="indirect"` respectively, and `br=FALSE` (default). If `br=TRUE`, the bias-reduced estimators $\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}$ and $\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}$ are returned instead. Argument `estim="Hill"` (default) calls the Hill estimator (function `mop` in the package `evt0`), and `estim="tindexp"` calls `tindexp`; setting `br=TRUE` calls the bias-reduced versions of these estimators. The choice of k_n is also an option through `k`, and by default $k_n = \widehat{k}_n^{\mathrm{H}}$ (if `estim="Hill"`) or $k_n = \widehat{k}_n^{\mathrm{E}}$ (if `estim="tindexp"`).
- Extreme quantile estimation: The estimator $\widetilde{q}_{\beta_n}^{\star}$ is computed in the function `extQuant` with argument `br=FALSE` (default). If `br=TRUE`, the bias-reduced estimator $\widetilde{q}_{\beta_n}^{\star,\mathrm{RB}}$ is returned instead. The other arguments are those of `extExpect`.
- Data sets: `austria`, `belgium`, `commerzbank`, `finland`, `france`, `greece`, `italy`, `namibia`, `netherlands`, `newzealand`, `secura` and `southafrica`, as described in Section 5.

4 Simulation study

We study the finite-sample performance of our estimators on simulated data in order to assess the importance of bias reduction in extreme expectile estimation. For that purpose and in order to get a good overview of practical performance, we consider the following heavy-tailed distributions for Y (see Table 1):

- 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 $y > 0$. The interesting point here is that the choices of γ and ρ are free, meaning that for a fixed γ we can make the second-order parameter ρ vary in order to generate scenarios with various degrees of difficulty in the estimation. We consider here $\rho = -5, -1, -0.5$, corresponding respectively to an easy, medium, and hard estimation problem.
- A Generalised Pareto Distribution (GPD) with tail index $\gamma > 0$ and unit scale, *i.e.* $\overline{F}(y) = (1 + \gamma y)^{-1/\gamma}$ for $y > 0$. Here $\rho = -\gamma$. For γ close to 0, ρ will then also be close to 0, meaning that we expect the estimation problem to be difficult when the data are generated from this distribution.

For each of these distributions (three Burr distributions and one GPD), we consider the cases $\gamma = 0.1, 0.2, 0.3, 0.4$. This gives 16 cases in total. In each case, we simulate $N = 1{,}000$ data sets $Y_1, \ldots, Y_n$ of $n = 1{,}000$ independent realisations of Y, with survival distribution function $\overline{F}$. We estimate the expectile of level $\beta_n = 1 - 5/n = 0.995$ using five methodologies:

(i) The bias-reduced extrapolated LAWS estimator $\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}$ with the expectile-based, bias-reduced tail index estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$,

(ii) The bias-reduced extrapolated indirect estimator $\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}$ with the expectile-based, bias-reduced tail index estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$,

(iii) The bias-reduced extrapolated LAWS estimator $\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}$ with the bias-reduced Hill estimator $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$,

(iv) The bias-reduced extrapolated indirect estimator $\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}$ with the bias-reduced Hill estimator $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$,

(v) (As a benchmark) The extrapolated LAWS estimator $\widehat{\xi}^{\star}_{\beta_n}$ (without bias reduction) with the bias-reduced Hill estimator $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$.

Comparing methods (i) and (iii) on the one hand, and (ii) and (iv) on the other hand, allows us to see the influence of the choice of tail index estimator. Comparing (iii), (iv) and (v) makes it possible to assess the benefit of the bias reduction method in the expectile extrapolation. To get a further idea of the difference between using the bias-reduced tail index estimator $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ introduced in the current paper and the bias-reduced Hill estimator $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$, we also record the values of $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ and $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$.

We assess finite-sample performance by computing the following quantities:

- The relative bias, variance and mean-squared error of the extrapolated expectile estimators. For the bias-reduced extrapolated LAWS estimator $\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}$, that is

$$\mathrm{RBias}(\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{\widehat{\xi}^{\star,\mathrm{RB},(j)}_{\beta_n}}{\xi_{\beta_n}} - 1\right) \text{ and } \mathrm{RMSE}(\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{\widehat{\xi}^{\star,\mathrm{RB},(j)}_{\beta_n}}{\xi_{\beta_n}} - 1\right)^2$$

  (where $\widehat{\xi}^{\star,\mathrm{RB},(j)}_{\beta_n}$ is calculated on the jth sample) and $\mathrm{RVar} = \mathrm{RMSE} - \mathrm{RBias}^2$ is the relative variance. Similarly for the indirect estimator $\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}$ and the non-bias-reduced estimator $\widehat{\xi}^{\star}_{\beta_n}$.

- The (classical) bias, variance and mean-squared error of the estimators $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ and $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$.

All these quantities are calculated for k_n chosen to be (in each sample) the selected value $\widehat{k}^{\mathrm{E}}_n$ or $\widehat{k}^{\mathrm{H}}_n$ as appropriate, depending on whether the estimator $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ or $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ is used, and they are reported in Tables C.1, C.2, C.3 and C.4 in the Appendix. We also record, and report, median expectile and tail index estimates over $k_n \in \{2, 3, \ldots, 450\}$ and the corresponding log-mean-squared errors in Figures C.1, C.2, C.3, C.4 and C.5 in this same Appendix.

We conclude from this simulation study that, on our tested cases, the bias reduction scheme is very effective: as a consequence, the RMSE of the bias-reduced estimates is often one and sometimes two orders of magnitude lower than the RMSE of the standard extrapolated estimates. This is true across a wide range of values of k_n, as shown in Figures C.1, C.2 and C.3, where it is seen that overall the bias-reduced estimators based on $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ seem to have an advantage in terms of bias, while those based on $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ have a lower MSE overall. There does not seem to be a clear winner between (bias-reduced) direct and indirect estimators. It also appears (from considering the case of the GPD distribution) that the bias-reduced estimator $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ is particularly interesting for values of ρ close to 0 both in terms of bias and MSE, and is competitive otherwise; note that for large $|\rho|$, the Burr distribution gets very close to the Pareto distribution for which the Hill estimator is the Maximum Likelihood estimator and known to be optimal, so it is not reasonable to expect that the estimator $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ would be more accurate than the bias-reduced Hill estimator in such cases. The estimator $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ appears to be a useful complement to available tail index estimation devices.

5 Applications on real data

We apply our methodology to three data sets, from insurance, economics, and finance, as a way to illustrate the applicability of our expectile and tail index estimators.

5.1 Reinsurance premium estimation

Reinsurance is a very important way of mitigating risk associated with high-impact events such as extreme climate episodes. Reinsurance contracts typically involve two insurance companies A and B; by the terms of the contract, company A transfers to company B (totally or partially) the risk associated to events involving large claims. Here we focus on the case when risk is totally transferred, which is also called excess-of-loss reinsurance. Under such a policy, when a claim occurs, company A pays the claim amount up to a certain amount R decided in the reinsurance contract, called retention level, and company B underwrites all losses above that amount R. In other words, if the total claim amount is Y, company A pays $\min(Y, R)$, and company B pays $\max(Y - R, 0)$.

A crucial task to decide the terms of a reinsurance contract is to accurately price this contract, which leads to the calculation of the so-called reinsurance premium. A first, natural approach to do this is to use the net premium principle (see *e.g.* Chapters 4 and 5 in Kaas et al. (2008)), namely

$$\Pi(R) = \mathbb{E}[\max(Y - R, 0)] = \int_R^\infty \overline{F}(x)dx,$$

where $\overline{F}$ is the survival function of Y. However, paying company B this net average premium would not protect that company B from a catastrophic loss much higher than its average value. A solution to this problem developed in the actuarial literature over the last 25 years has been to consider more conservative premium principles, including the distorted premiums introduced in Wang (1996):

$$\Pi_g(R) = \int_R^\infty g(\overline{F}(x))dx,$$

where $g : [0, 1] \to [0, 1]$ is a nondecreasing concave function such that $g(0) = 0$ and $g(1) = 1$, called the distortion function. The choice $g(x) = x$ leads to the net premium principle, but there are several reasonable ways of choosing a function g leading to a more conservative (*i.e.* higher) premium, such as the Dual Power function, or the Proportional Hazards function (we refer to, among others, the paper of Wang (1995) and Chapter 3 of Dickson (2016)).

Since in reinsurance the retention level R should be considered as a high (and therefore rarely observed) level of claim amount, the calculation of the reinsurance premium is very closely linked to the right tail of the distribution of the claim amount Y. This motivated Vandewalle and Beirlant (2006) to develop an extreme value theory-based method for the estimation of $\Pi_g(R)$ (a different, somewhat linked theory for Wang distortion risk measures conditionally on the loss being high is provided in El Methni and Stupfler (2017)). In particular, they proved that if Y is heavy-tailed with tail index γ and $g(1/\cdot)$ is regularly varying with tail index $\delta < -\gamma$, *i.e.* $g(1/(ty))/g(1/t) \to y^\delta$ as $t \to \infty$, where $\delta < -\gamma < 0$, then the following limiting relationship holds between the distorted premium, the retention level and the probability of exceeding that level:

$$\lim_{R \to +\infty} \frac{\Pi_g(R)}{R\, g(\overline{F}(R))} = \frac{1}{-\delta\gamma^{-1} - 1}.$$

An interesting version of that relationship is found when $R = \xi_\beta$ for $\beta \uparrow 1$. In that case, using Equation (11), we get

$$\lim_{\beta \uparrow 1} \frac{\Pi_g(\xi_\beta)}{\xi_\beta\, g(1 - \beta)} = \frac{(\gamma^{-1} - 1)^{-\delta}}{-\delta\gamma^{-1} - 1}. \tag{17}$$

In the case of the net premium, $g(x) = x$, so $\delta = -1$ and we find $\Pi_g(\xi_\beta) = \Pi(\xi_\beta) \sim (1 - \beta)\xi_\beta$; in this asymptotic equivalence, the tail index γ does not appear anymore, and the expectile is in some sense an asymptotic inverse of the function $R \mapsto \Pi(R)/R$ that represents the proportion of the retention level R paid on average per claim by company B (in other words, if $\Pi(R)/R = \pi$, then company B contributes πR to the payment of the average claim).

Given this relevance of expectiles to premium calculation for large claims, we propose to estimate the reinsurance premium $\Pi_g(R)$ (for a large retention level R) using Equation (17) and our bias-reduced extreme expectile estimation methodology. For that purpose, we consider the well-known Secura Belgian Re data used in Vandewalle and Beirlant (2006), available in our package `Expectrem` as well as several other R packages such as `ReIns`, `CASdatasets` and `ltmix`. This data set contains $n = 370$ inflation-adjusted automobile claim amounts (from 1988 to 2001), larger than €1,200,000. We consider two premium principles: the net premium ($g(x) = x$) and Dual Power ($g(x) = 1 - (1-x)^{\kappa}$) principles, and to allow us to compare our results with those of Vandewalle and Beirlant (2006), we take $\kappa = 1.366$. This choice was already recommended in Wang (1996). We estimate the associated premiums $\Pi_g(R)$ with the statistic

$$\widehat{\Pi}_g^{\star}(\xi_\beta) = \frac{(\overline{\gamma}^{-1} - 1)^{-\delta}}{-\delta\overline{\gamma}^{-1} - 1} \widehat{\xi}_\beta^{\star,\mathrm{RB}}\, g(1-\beta).$$

Here $\widehat{\xi}_\beta^{\star,\mathrm{RB}}$ denotes the bias-reduced LAWS extrapolated estimator calculated using either our bias-reduced asymptotic proportionality tail index estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ or its bias-reduced Hill counterpart $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$, and $\overline{\gamma}$ is taken to be the same tail index estimator as the one used in the extrapolation. We also compare this estimator with the one in which the bias-reduced estimator $\overline{\xi}_\beta^{\star,\mathrm{RB}}$ is replaced by the standard, non-bias-reduced extrapolated LAWS estimator calculated with $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ (and thus $\overline{\gamma} = \widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ everywhere). The estimated reinsurance premiums are represented in Figure 2 on a fine grid of values of β_n; this yields curves of estimates of $\Pi_g(R)$ in the tail region $R \to \infty$, that we compare to the premium curve obtained by Vandewalle and Beirlant (2006). Indirect estimators are here almost identical to their LAWS counterparts, and are not reported for the sake of readability.

We draw two conclusions from Figure 2. First, our bias-reduced expectile-based estimators constructed from Formula (17) are at first quite conservative, but become very close to those of Vandewalle and Beirlant (2006) when the retention level is large, confirming the accuracy of our (bias-reduced) extreme expectile estimators. Interestingly, the point estimate based on our proposed tail index estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ is slightly less conservative than the others for the largest values of R that we consider. While policymakers and regulators would favour higher (*i.e.* more pessimistic) estimates such as those given by Vandewalle and Beirlant (2006), more optimistic (*i.e.* lower) assessments of risk may be interesting for insurance companies, because lower premiums paid by consumers translate into improved competitivity on insurance markets. Second, it is clearly seen that without the bias reduction scheme that we propose, the expectile-based estimates seem to be very poor and a long way off the other estimates we consider. This example therefore clearly emphasises the importance of the bias reduction methodology proposed in this paper as far as expectile-based estimation is concerned.

5.2 Approximation of the Gini index

We now showcase how our methodology can be applied in economics through the example of the estimation of the Gini index. This economic indicator measures the statistical dispersion (and therefore inequality) of income within a country: the Gini index of a country with n workers having respective incomes $Y_1, \ldots, Y_n$ is given by

$$\overline{G} = \frac{\sum_{i=1}^n \sum_{j=1}^n |Y_i - Y_j|}{2n \sum_{i=1}^n Y_i}.$$

A higher Gini index means higher inequality of income within the sampled population. Of course, in practice n is very large (of the order of millions, if not a billion) and income data is typically very sensitive, so to estimate the Gini index of a country using the above formula, it is generally the case that a representative survey of incomes representing all the categories of workers is carried out. Ensuring representativity in this context can be extremely difficult and time- and labour-intensive. In particular, it is the case that substantial left-censoring or left-truncation can be present, as it is

reasonable to imagine that accurately sampling from the lowest-paid workers is hard, for example because of job unstability, or labour market law violations from employers including minimum wage underpayment or illegal employment of foreign workers. Accurately representing the left tail of the income distribution, which is key if the above definition of the Gini index is to be used, can therefore be a tall order. An alternative solution putting more weight on the right tail is to model the distribution of income within a country by a heavy-tailed distribution (as done for example in the recent work by Gardes and Girard (2021)). This kind of approach has a long history in labour economics, see for instance Singh and Maddala (1976) and McDonald (1984). A particularly interesting model uses the Burr distribution with parameters γ (the tail index) and ρ (the second-order parameter). It can then be shown that the Gini index should be

$$G = G(\gamma, \rho) = 1 - \frac{\Gamma(-1/\rho)\Gamma((\gamma-2)/\rho)}{\Gamma(-2/\rho)\Gamma((\gamma-1)/\rho)}. \tag{18}$$

Here $\Gamma(\cdot)$ is Euler's Gamma function; see *e.g.* Chotikapanich and Griffiths (2000). This way of modelling the Gini index has the advantage to use only the right tail of the distribution of income, and is thus more robust to sampling inaccuracies in the left tail.

We propose here to estimate the Gini index for several countries using Formula (18) and the bias-reduced tail index estimates, and to compare our results with official Gini indices calculated by the CIA[1], World Bank[2] and Eurostat[3]. The countries and data considered are the following:

- The synthetic Eurostat[4] data set of incomes for Austria ($n = 5{,}977$), Belgium ($n = 6{,}159$), France ($n = 11{,}131$), Finland ($n = 11{,}370$), Greece ($n = 7{,}439$) and the Netherlands ($n = 10{,}131$).
- A data set of $n = 8{,}156$ Italian incomes for the year 2014, available from the Bank of Italy[5].
- A survey of $n = 9{,}656$ wages in Namibia during the period 2009-2010, provided by the Namibia Statistics Agency[6].
- A synthetic data set of $n = 11{,}315$ incomes in New Zealand during the year 2003, collected by the official Statistics New Zealand agency[7].
- The Filipino Family Income and Expenditure data set[8] of $n = 41{,}544$ incomes measured in the Philippines in 2015 by the Philippine Statistics Authority.
- The Living Conditions Survey 2014-2015[9] in South Africa, containing $n = 19{,}286$ wages.

The Gini coefficient is estimated with

$$\widehat{G}\left(\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}\right) = G\left(\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}, \overline{\rho}\right) \text{ and } \widehat{G}\left(\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}\right) = G\left(\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}, \overline{\rho}\right),$$

where the parametric form $G(\gamma, \rho)$ of the Gini index is defined in (18) and $\overline{\rho}$ is the second-order parameter estimator we use throughout in our bias reduction scheme (see Appendix A). Our estimates are reported in Table 2, where they are compared with official Gini indices calculated by the World Bank, CIA and/or Eurostat as appropriate, as well as with their versions $\widehat{G}\left(\widehat{\gamma}_{k_n}^{\mathrm{H}}\right)$ and $\widehat{G}\left(\widehat{\gamma}_{k_n}^{\mathrm{E}}\right)$ that

[1] https://www.cia.gov/library/publications/the-world-factbook/rankorder/2172rank.html
[2] https://data.worldbank.org/indicator/SI.POV.GINI
[3] https://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tessi190&plugin=1
[4] https://ec.europa.eu/eurostat/web/microdata/statistics-on-income-and-living-conditions
[5] https://www.bancaditalia.it/statistiche/tematiche/indagini-famiglie-imprese/bilanci-famiglie/distribuzione-microdati/index.html?com.dotmarketing.htmlpage.language=1
[6] https://nsa.org.na/microdata1/index.php/catalog/6
[7] https://www.stats.govt.nz/services/customised-data-services/statistics-for-university-staff-and-students/new-zealand-income-survey-super-surf
[8] https://www.kaggle.com/grosvenpaul/family-income-and-expenditure
[9] https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/608

do not feature any bias reduction. The first conclusion we can draw is that the bias reduction scheme applied to the asymptotic proportionality tail index estimator is very effective: the non-bias-reduced estimates are typically far from their bias-reduced counterparts as well as from official estimates, and when the non-bias-reduced estimate is sensible (in the example of the Philippines and South Africa), the bias-reduced estimate is not unreasonable either. The second conclusion is that the bias-reduced estimates based on $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ are competitive and sometimes even closer to the average Gini index (computed using the average Gini indices from our official sources) than the estimate using the bias-reduced Hill estimator. This shows that our bias-reduced asymptotic proportionality tail index estimator is a valuable resource in practical applications.

5.3 An analysis of financial returns

Our final real data example focuses on financial returns. The fact that expectiles can, in financial contexts, be interpreted in terms of the gain-loss ratio, makes them interesting in portfolio management. Recently Bellini and Di Bernardino (2017) have shown the practical interest of estimating extreme expectiles of series of financial log-returns series; in particular, they recommend the estimation of the expectile of level $\beta = 0.99855$, following their observation that it coincides with the quantile of level $\beta' = 0.99$ in the standard Gaussian case. In this example, we consider the series of the daily negative log-returns of the Commerzbank stock prices on the DAX30 stock exchange between March 6, 2012 and July 28, 2016, resulting in a sample $Y_1, \ldots, Y_n$ of size $n = 1{,}048$ plotted in the top left panel of Figure 3. To reduce the serial dependence in the observations, we filter our time series using a GARCH$(1,1)$ model:

$$Y_t = \sigma_t \varepsilon_t, \text{ with } \sigma_t > 0 \text{ such that } \sigma_t^2 = \sigma_t^2(Y_{t-1}^2, \sigma_{t-1}^2) = \mathfrak{c} + \mathfrak{a} Y_{t-1}^2 + \mathfrak{b}\sigma_{t-1}^2.$$

Here $\mathfrak{a}, \mathfrak{b}, \mathfrak{c} > 0$ are unknown coefficients, and (ε_t) is an unobserved independent nonconstant white noise sequence, *i.e.* such that $\mathbb{E}(\varepsilon) = 0$, $\mathbb{E}(\varepsilon^2) = 1$ and $\mathbb{P}(\varepsilon^2 = 1) < 1$. Under suitable conditions, this model has a stationary, nonanticipative solution, see Theorem 2.4 p.30 of Francq and Zakoïan (2010); this is in particular the case if $\mathfrak{a} + \mathfrak{b} < 1$. In this case the process σ_t is a function of $Y_{t-1}, Y_{t-2}, \ldots$ only. By positive homogeneity of expectiles, the conditional expectile for the next day given our observations up to time n is then

$$\xi_\beta(Y_{n+1} | Y_t, t \le n) = \sigma_{n+1}\, \xi_\beta(\varepsilon).$$

We estimate here an extreme conditional expectile for tomorrow given our knowledge of today, that is, the quantity $\xi_{\beta_n}(Y_{n+1}|Y_t, t \le n)$ with $\beta_n = \beta = 0.99855$. We first estimate $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ by Gaussian quasi-maximum likelihood (see Chapter 7 in Francq and Zakoïan (2010)) using the function `garch` in the R package `tseries`. This gives residuals $\widehat{\varepsilon}_i$ from the model, which we treat as independent and identically distributed copies of ε for the estimation of the tail index γ of ε, and $\xi_{\beta_n}(\varepsilon)$. Evidence that ε is indeed heavy-tailed is gathered in the two bottom panels of Figure 3 using exponential QQ-plots of the log-spacings (the somewhat erratic behaviour of the four points at the right end of the plot is quite typical because the very top observations have a large variance; see Ghosh and Resnick (2010) in the linked context of mean excess plots). Estimated tail indices of the residuals are the very similar $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}} \approx 0.253$ and $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}} \approx 0.270$ (selected k_n values are $\widehat{k}_n^{\mathrm{H}} = 50$ and $\widehat{k}_n^{\mathrm{E}} = 30$ respectively). The `garch` routine also provides an estimate $\widehat{\sigma}_n$ of the conditional standard deviation on day $n+1$, produced by manually inputting an initial value for the volatility on day 1 and iterating using the formula

$$\widehat{\sigma}_t^2 = \widehat{\sigma}_t^2(Y_{t-1}^2, \widehat{\sigma}_{t-1}^2) = \widehat{\mathfrak{c}} + \widehat{\mathfrak{a}} Y_{t-1}^2 + \widehat{\mathfrak{b}}\widehat{\sigma}_{t-1}^2, \text{ for } t \ge 2.$$

This eventually yields the conditional extreme expectile estimates

$$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}(Y_{n+1}|Y_t, t \le n) = \widehat{\sigma}_{n+1}\, \widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}(\varepsilon) \text{ and } \widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}(Y_{n+1}|Y_t, t \le n) = \widehat{\sigma}_{n+1}\, \widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}(\varepsilon).$$

We represent these estimates in the top right panel of Figure 3 as a function of the observation on day n; in other words, we let $Y_n = y_n$ vary in $\widehat{\sigma}_{n+1}^2(y_n^2, \widehat{\sigma}_n^2)$, as a way of evaluating the influence

of the nth observation on the dynamic extreme expectile prediction for the next day. Our estimates $\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}(Y_{n+1}|Y_t, t \leq n)$ and $\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}(Y_{n+1}|Y_t, t \leq n)$ are calculated using either $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ or $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$, and are compared to their counterpart using the standard, non-bias-reduced Weissman estimate $\widehat{\xi}^{\star}_{\beta_n}(Y_{n+1}|Y_t, t \leq n)$ extrapolated with $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$. They are also compared with dynamic bias-reduced Weissman quantile estimates of level $\beta'_n = \beta' = 0.99$, calculated using either $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ or $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$:

$$\widetilde{q}^{\star,\mathrm{RB}}_{\beta'_n}(Y_{n+1}|Y_t, t \leq n) = \widehat{\sigma}_{n+1}\, \widetilde{q}^{\star,\mathrm{RB}}_{\beta'_n}(\varepsilon),$$

where the expression of $\widetilde{q}^{\star,\mathrm{RB}}_{\beta'_n}(\varepsilon)$ (adapted here with the use of residuals) is given in Section 3.3.

The expectile estimates give similar results, thus suggesting that bias does not play an important role in this example. Estimated expectiles are substantially larger than estimated quantiles; this justifies further the non-Gaussian behaviour of the returns, and also means 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.

Acknowledgements

This research is supported by the French National Research Agency under the grant ANR-19-CE40-0013/ExtremReg project. S. Girard also acknowledges the support of the Chair Stress Test, Risk Management and Financial Steering, led by the French Ecole Polytechnique and its Foundation and sponsored by BNP Paribas, as well as the support of the French National Research Agency in the framework of the Investissements d'Avenir program (ANR-15-IDEX-02). G. Stupfler also acknowledges support from an AXA Research Fund Award on "Mitigating risk in the wake of the COVID-19 pandemic".

References

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

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

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

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

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

Cai, J.-J., de Haan, L., and Zhou, C. (2013). Bias correction in extreme value statistics with index around zero. *Extremes*, 16(2):173–201.

Chotikapanich, D. and Griffiths, W. E. (2000). Posterior distributions for the Gini coefficient using grouped data. *Australian & New Zealand Journal of Statistics*, 42(4):383–392.

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

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

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

Daouia, A., Girard, S., and Stupfler, G. (2021). ExpectHill estimation, extreme risk and heavy tails. *Journal of Econometrics*, to appear, https://doi.org/10.1016/j.jeconom.2020.02.003.

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

Dickson, D. C. (2016). *Insurance risk and ruin*. Cambridge University Press.

Ehm, W., Gneiting, T., Jordan, A., and Krüger, F. (2016). Of quantiles and expectiles: consistent scoring functions, choquet representations and forecast rankings. *Journal of the Royal Statistical Society: Series B*, 78:505–562.

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

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer.

Fraga Alves, M. I., Gomes, M. I., and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, 60(2):193–214.

Francq, C. and Zakoïan, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. John Wiley & Sons.

Gardes, L. and Girard, S. (2021). On the estimation of the variability in the distribution tail. *Test*, to appear, https://hal.inria.fr/hal-02400320.

Ghosh, S. and Resnick, S. (2010). A discussion on mean excess plots. *Stochastic Processes and their Applications*, 120:1492–1517.

Girard, S., Stupfler, G., and Usseglio-Carleve, A. (2020a). Extreme conditional expectile estimation in heavy-tailed heteroscedastic regression models. Available at https://hal.inria.fr/hal-02531027.

Girard, S., Stupfler, G., and Usseglio-Carleve, A. (2020b). Nonparametric extreme conditional expectile estimation. *Scandinavian Journal of Statistics*, to appear, https://doi.org/10.1111/sjos.12502.

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

Goegebeur, Y., Beirlant, J., and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference*, 140(9):2632–2652.

Gomes, M., Pestana, D., and Caeiro, F. (2009). A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Statistics and Probability Letters*, 79(3):295–303.

Gomes, M. I., Brilhante, M. F., and Pestana, D. (2016). New reduced-bias estimators of a positive extreme value index. *Communications in Statistics-Simulation and Computation*, 45(3):833–862.

Gomes, M. I. and Martins, M. J. (2002). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes*, 5(1):5–31.

Gomes, M. I. and Pestana, D. (2007). A sturdy reduced-bias extreme quantile (VaR) estimator. *Journal of the American Statistical Association*, 102(477):280–292.

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

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

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

Hua, L. and Joe, H. (2011). Second order regular variation and conditional tail expectation of multiple risks. *Insurance: Mathematics and Economics*, 49:537–546.

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

Kaas, R., Goovaerts, M., Dhaene, J., and Denuit, M. (2008). *Modern actuarial risk theory: using R*, volume 128. Springer Science & Business Media.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46:33–50.

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

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52(3):647–665.

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

Padoan, S. and Stupfler, G. (2020). Extreme expectile estimation for heavy-tailed time series. Available at https://arxiv.org/abs/2004.04078.

Resnick, S. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer.

Singh, S. K. and Maddala, G. S. (1976). A function for size distribution of incomes. *Econometrica*, 44(5):963–970.

Sobotka, F. and Kneib, T. (2012). Geoadditive expectile regression. *Computational Statistics & Data Analysis*, 56:755–767.

Vandewalle, B. and Beirlant, J. (2006). On univariate extreme value statistics and the estimation of reinsurance premiums. *Insurance: Mathematics and Economics*, 38(3):441–459.

Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, 17(1):43–54.

Wang, S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, 26(1):71–92.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815.

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

Distribution (parameters)	Density function	γ	ρ	b
Pareto ($\alpha > 0$)	$\alpha t^{-\alpha-1}$ $(t > 1)$	$1/\alpha$	$-\infty$	0 (by convention)
Generalised Pareto ($\sigma, \xi > 0$)	$\sigma^{-1}\left(1+\xi t/\sigma\right)^{-1-1/\xi}$ $(t > 0)$	ξ	$-\xi$	1
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 ($\alpha, \beta > 0$)	$\alpha\beta t^{\alpha-1}\left(1+t^{\alpha}\right)^{-\beta-1}$ $(t > 0)$	$1/(\alpha\beta)$	$-1/\beta$	1
Fréchet ($\alpha > 0$)	$\alpha t^{-\alpha-1}\exp\left(-t^{-\alpha}\right)$ $(t > 0)$	$1/\alpha$	-1	$1/2$
Fisher ($\nu_1, \nu_2 > 0$)	$\dfrac{(\nu_1/\nu_2)^{\nu_1/2}}{B(\nu_1/2, \nu_2/2)} t^{\nu_1/2-1}(1+\nu_1 t/\nu_2)^{-(\nu_1+\nu_2)/2}$ $(t > 0)$	$2/\nu_2$	$-2/\nu_2$	$\dfrac{\nu_1+\nu_2}{\nu_2+2}[B(\nu_1/2, \nu_2/2)]^{2/\nu_2}(\nu_2/2)^{2/\nu_2}$
Inverse-Γ ($\alpha, \beta > 0$)	$\dfrac{\beta^{\alpha}}{\Gamma(\alpha)} t^{-\alpha-1}\exp(-\beta/t)$ $(t > 0)$	$1/\alpha$	$-1/\alpha$	$[\Gamma(\alpha+1)]^{1/\alpha}/(\alpha+1)$
Cauchy ($c > 0$)	$\dfrac{c}{\pi(c^2+t^2)}$	1	-2	$2\pi^2/3$
Student ($\nu > 0$)	$\dfrac{1}{\sqrt{\nu\pi}}\dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\left(1+\dfrac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	$1/\nu$	$-2/\nu$	$\dfrac{\nu(\nu+1)}{\nu+2}\left(\dfrac{\Gamma\left(\frac{\nu+1}{2}\right)\nu^{(\nu-1)/2}}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\right)^{-2/\nu}$

Table 1: 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 Hall-Weiss distribution is taken from Example 1 in Hua and Joe (2011).

Country	$\widehat{G}(\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{G}(\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	World Bank	CIA	Eurostat	Average Gini
Austria	28.4 (30)	28.1 (38.1)	30.8 (2013)	30.5 (2015)	27 (2013)	29.4
Belgium	26.2 (27.7)	27.5 (37.8)	27.7 (2013)	25.9 (2013)	25.9 (2013)	26.5
Finland	26.4 (27.5)	27.7 (37.5)	27.2 (2013)	27.2 (2016)	25.4 (2013)	26.6
France	36 (37.6)	35.3 (46.5)	32.5 (2013)	29.3 (2016)	30.1 (2013)	30.6
Greece	35.2 (37)	35.8 (45.9)	36.1 (2013)	36.7 (2012)	34.4 (2013)	35.7
Italy	31.8 (33.3)	32.1 (43.9)	34.7 (2014)	31.9 (2012)	32.8 (2013)	33.1
Namibia	60.6 (62.9)	59.9 (64.5)	61 (2009)	59.7 (2010)	N/A	60.4
Netherlands	25.2 (26.4)	27.1 (37)	28.1 (2013)	30.3 (2015)	25.1 (2013)	27.8
New Zealand	35.6 (37)	34.2 (45.9)	N/A	36.2 (1997)	N/A	36.2
Philippines	38.7 (39.5)	38.4 (46.5)	44.4 (2015)	44.4 (2015)	N/A	44.4
South Africa	58.6 (60.5)	60.1 (64.4)	63 (2014)	62.5 (2013)	N/A	62.8

Table 2: Estimated Gini indices per country using the estimators $\widehat{G}(\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$ (reported along with $\widehat{G}(\widehat{\gamma}_{k_n}^{\mathrm{H}})$ between brackets) and $\widehat{G}(\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$ (reported along with $\widehat{G}(\widehat{\gamma}_{k_n}^{\mathrm{H}})$ between brackets), and official Gini indices (reported along with year of publication). The last column gives the average Gini index calculated from official estimations. Figures highlighted in violet indicate the Burr model-based estimate closest to the average Gini coefficient calculated using official figures.

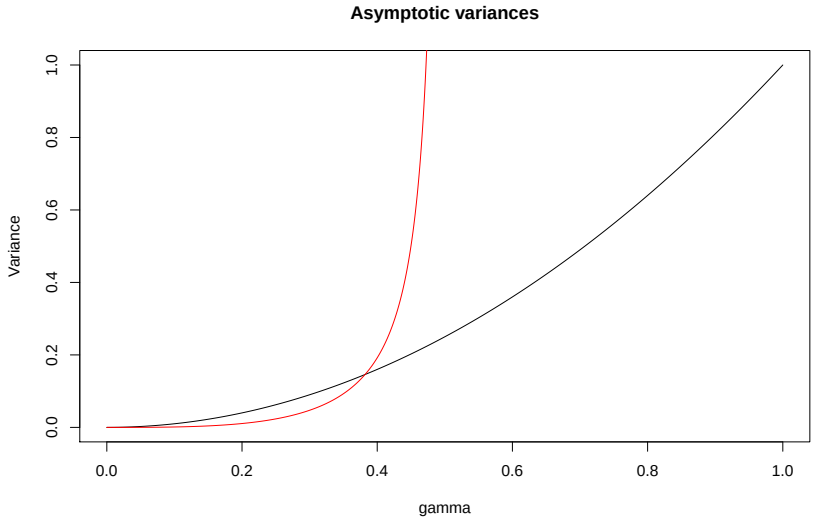

Figure 1: Asymptotic variances of $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ (black curve) and $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ (red curve) as functions of $\gamma \in (0, 1/2)$.

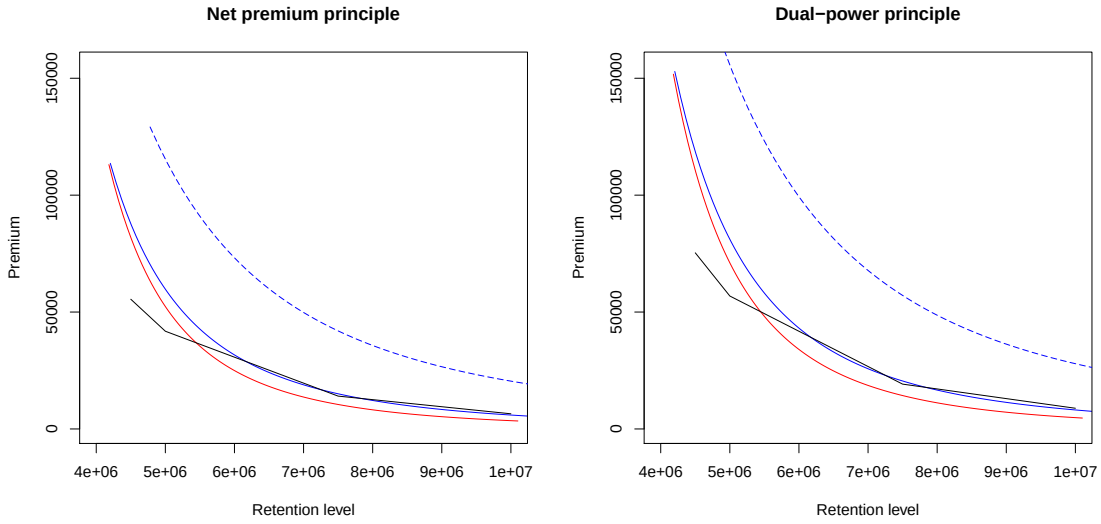

Figure 2: Secura Belgian Re insurance data, estimated distorted premium $\Pi_g(R)$ as function of the retention level $R = \xi_\beta$ for β ranging from $1-10/n \approx 0.973$ to $1-1/(8n) \approx 0.9997$ (here $n = 370$). The premiums are estimated using $\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}$ and $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ (solid blue curve), $\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}$ and $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ (solid red curve) and $\widehat{\xi}^{\star}_{\beta_n}$ and $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ (dotted blue curve). The black curve is constructed by linear interpolation using the estimates found in Vandewalle and Beirlant (2006). Left panel: net premium principle ($g(x) = x$). Right panel: Dual Power principle ($g(x) = 1-(1-x)^\kappa$) with $\kappa = 1.366$.

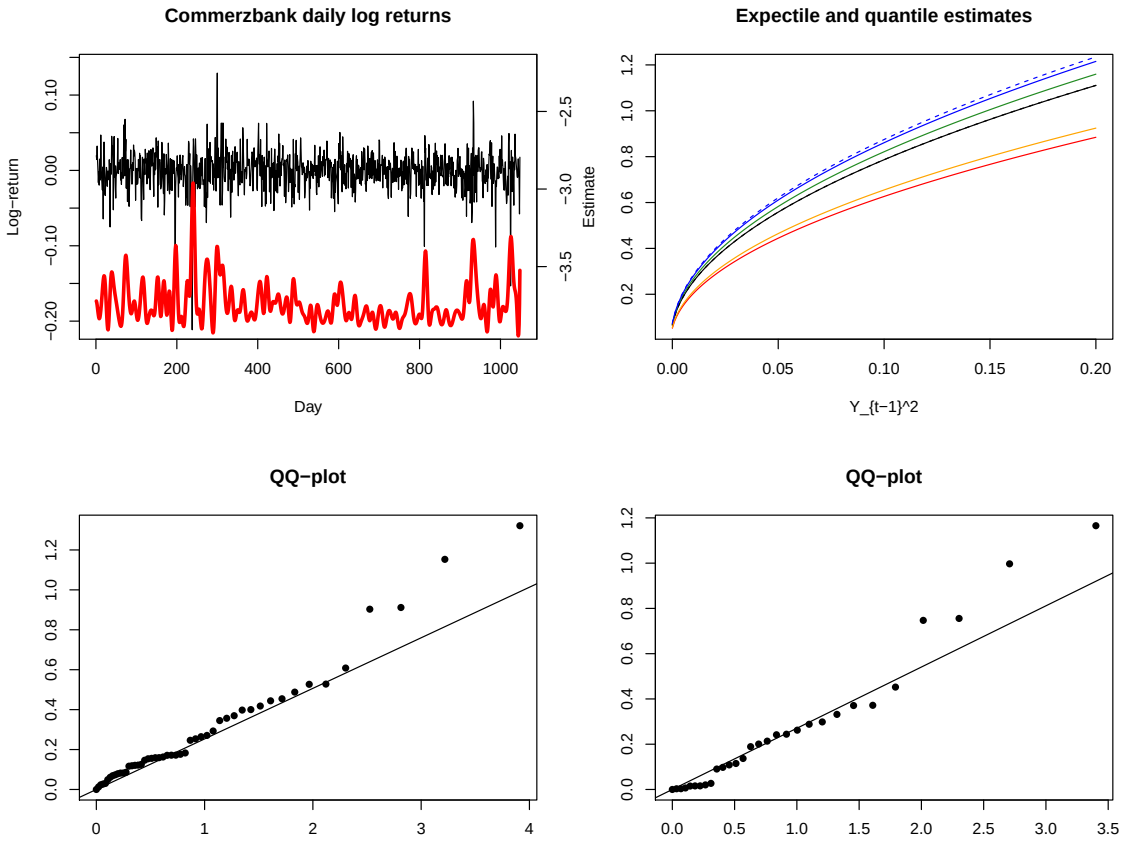

Figure 3: Commerzbank daily log-returns data between March 6, 2012 and July 28, 2016, sample size $n = 1{,}048$. In the top left panel, daily (positive) log-returns (black curve) and GARCH log-volatility estimates $t \mapsto \log \widehat{\sigma}_t$ (red bold curve, smoothed using the R function `smooth.spline` with smoothing parameter $\lambda = 0$). In the top right panel, dynamic extreme expectile (level $\beta_n = 0.99855 \approx 1-2/n$) and quantile (level $\beta_n' = 0.99$) estimates of Y_{n+1} given past observations as functions of $Y_n = y_n$: estimates $\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}(Y_{n+1}|Y_t, t \le n)$ and $\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}(Y_{n+1}|Y_t, t \le n)$ computed with $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ (respectively solid and dotted black curves) and corresponding estimates computed with $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ (respectively solid and dotted blue curves), estimate $\widehat{\xi}^{\star}_{\beta_n}(Y_{n+1}|Y_t, t \le n)$ computed with $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ (green curve), estimate $\widetilde{q}^{\star,\mathrm{RB}}_{\beta_n'}(Y_{n+1}|Y_t, t \le n)$ computed with $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n}$ (red curve) and $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}$ (orange curve). Bottom left (respectively bottom right): Exponential QQ-plot of the log-spacings $\log(\widehat{\varepsilon}_{n-i+1,n}/\widehat{\varepsilon}_{n-k^*,n})$, $1 \le i \le k^* = 50$ (respectively $1 \le i \le k^* = 30$). The straight line has slope $\widehat{\gamma}^{\mathrm{H,RB}}_{k_n} = 0.253$ (respectively $\widehat{\gamma}^{\mathrm{E,RB}}_{k_n} = 0.270$).

Supplementary Material for "On automatic bias reduction for extreme expectile estimation"

Stéphane Girard, Gilles Stupfler & Antoine Usseglio-Carleve

This supplementary material contains extra details on the estimators of the second-order extreme value parameters b and ρ that we use, all necessary mathematical proofs, and a complete set of numerical results from our simulation study.

A Estimation of second-order extreme value parameters

We explain here how we estimate the second-order parameters ρ and b which are ubiquitous in our bias reduction procedures. The estimators we use are introduced in Gomes and Martins (2002) and Fraga Alves et al. (2003). We start by the estimation of ρ. Let

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

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

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

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

$$\widehat{\rho}_{\kappa_n}^{(\tau)} = -\left|\frac{3(T_{\kappa_n}^{(\tau)}-1)}{T_{\kappa_n}^{(\tau)}-3}\right|. \tag{19}$$

This estimator is implemented in the R function `mop` from the package `evt0`. In this package, $\kappa_n = \lfloor n^{0.999} \rfloor$, 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 (see Section 3.2 in Gomes et al. (2016) for more details). According to Proposition 2.1 in Caeiro et al. (2005), these choices ensure, if ρ is large enough (a calculation analogue to that of Remark 2.2 in Caeiro et al. (2005) provides roughly $\rho > -249.75$, which will cover all practical applications), that $(\widehat{\rho}_{\kappa_n}^{(\tau)} - \rho)\log(n) = o_{\mathbb{P}}(1)$ as required in our asymptotic results. An estimator of b is then

$$\widehat{b}_{\kappa_n} = \left(\frac{\kappa_n}{n}\right)^{\overline{\rho}} \frac{\left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-\overline{\rho}}\right)\left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n} U_i\right) - \left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-\overline{\rho}} U_i\right)}{\left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-\overline{\rho}}\right)\left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-\overline{\rho}} U_i\right) - \left(\frac{1}{\kappa_n}\sum_{i=1}^{\kappa_n}\left(\frac{i}{\kappa_n}\right)^{-2\overline{\rho}} U_i\right)}, \tag{20}$$

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).

B Mathematical proofs

The proof of Theorem 1 uses the following auxiliary result, somewhat similar to results in Girard et al. (2020b) and Padoan and Stupfler (2020). We provide a very concise proof for the sake of completeness.

Proposition 1. *Suppose that $\mathbb{E}|Y_-|^{2+\delta} < \infty$ for some $\delta > 0$. Assume further that $\mathcal{C}_2(\gamma, \rho, A)$ holds with $0 < \gamma < 1/2$, and let k_n be such that $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. If $\sqrt{k_n} A(n/k_n) \to \lambda_1 \in \mathbb{R}$ and $\sqrt{k_n}/q(1-k_n/n) \to \lambda_2 \in \mathbb{R}$, then*

$$\sqrt{k_n}\left(\widehat{\gamma}^{\mathrm{E}}_{k_n} - \gamma\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\frac{\gamma(\gamma^{-1}-1)^{1-\rho}}{1-\gamma-\rho}\lambda_1 + \gamma^2(\gamma^{-1}-1)^{\gamma+1}\mathbb{E}[Y]\lambda_2, \frac{\gamma^3(1-\gamma)}{1-2\gamma}\right).$$

Proof of Proposition 1. Let $z \in \mathbb{R}$ be arbitrary and focus on the probability

$$\Phi_n(z) = \mathbb{P}\left(\sqrt{k_n}\left(\frac{n\widehat{\overline{F}}_n(\widehat{\xi}_{1-k_n/n})}{k_n} - (\gamma^{-1}-1)\right) \le z\right) = \mathbb{P}\left(\widehat{\overline{F}}_n(\widehat{\xi}_{1-k_n/n}) \le (1-\alpha_n)(\theta + z/\sqrt{k_n})\right)$$

where on the right-hand side, we let $\alpha_n = 1-k_n/n$ and $\theta = \gamma^{-1}-1$. Equivalently $\Phi_n(z) = \mathbb{P}(\widehat{\xi}_{1-k_n/n} \ge \widehat{q}_{\alpha'_n})$, where $\alpha'_n = 1-(1-\alpha_n)(\theta + z/\sqrt{k_n})$, and so

$$\Phi_n(z) = \mathbb{P}\left(\sqrt{k_n}\left(\frac{\widehat{\xi}_{1-k_n/n}}{\xi_{1-k_n/n}} - 1\right) \ge \sqrt{k_n}\left(\frac{\widehat{q}_{\alpha'_n}}{q_{\alpha'_n}} - 1\right)\frac{q_{\alpha'_n}}{\xi_{1-k_n/n}} + \sqrt{k_n}\left(\frac{q_{\alpha'_n}}{\xi_{1-k_n/n}} - 1\right)\right).$$

Setting

$$\Theta_n = \sqrt{k_n}\left(\frac{\widehat{\xi}_{1-k_n/n}}{\xi_{1-k_n/n}} - 1\right),\ \Theta'_n = \Theta'_n(z) = \sqrt{k_n}\left(\frac{\widehat{q}_{\alpha'_n}}{q_{\alpha'_n}} - 1\right) \text{ and } \delta_n = \delta_n(z) = \frac{q_{\alpha'_n}}{\xi_{1-k_n/n}} - 1,$$

we find

$$\Phi_n(z) = \mathbb{P}(\Theta_n \ge \Theta'_n(1+\delta_n) + \sqrt{k_n}\,\delta_n) = \mathbb{P}(\Theta_n - \Theta'_n + \mathrm{o}_{\mathbb{P}}(1) \ge \sqrt{k_n}\,\delta_n)$$

by Proposition 1 in Daouia et al. (2020) and the consistency of $\widehat{q}$ at the intermediate level α'_n. The conclusion follows from finding the limit of $\Phi_n(z)$ thanks to an application of Proposition 1 and Theorem 1 in Daouia et al. (2020) and straightforward calculations. □

Proof of Theorem 1. We prove the result when the intermediate and extrapolated expectile estimators in the bias reduction method are the LAWS versions; the proof for the indirect estimators is identical. Note the identity

$$\frac{1}{\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}} = 1 + \left(\frac{1}{\widehat{\gamma}^{\mathrm{E}}_{k_n}} - 1\right) \times \frac{1}{1+\overline{r}(1-k_n/n)}.$$

It follows that

$$\frac{1}{\widehat{\gamma}^{\mathrm{E,RB}}_{k_n}} - \frac{1}{\gamma} = \left(\frac{1}{\widehat{\gamma}^{\mathrm{E}}_{k_n}} - \frac{1}{\gamma}\right) \times \frac{1}{1+\overline{r}(1-k_n/n)} - \left(\frac{1}{\gamma} - 1\right) \times \frac{\overline{r}(1-k_n/n)}{1+\overline{r}(1-k_n/n)}. \tag{21}$$

By Proposition 1,

$$\sqrt{k_n}\left(\frac{1}{\widehat{\gamma}^{\mathrm{E}}_{k_n}} - \frac{1}{\gamma}\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(-\frac{(\gamma^{-1}-1)^{1-\rho}}{\gamma(1-\gamma-\rho)}\lambda_1 - (\gamma^{-1}-1)^{\gamma+1}\mathbb{E}[Y]\lambda_2, \frac{1-\gamma}{\gamma(1-2\gamma)}\right). \tag{22}$$

It follows from our assumptions on $\overline{\gamma}$, $\overline{b}$ and $\overline{\rho}$, combined with the law of large numbers, the Gaussian approximation of the tail empirical distribution function in Theorem 5.1.4 p.161 in de Haan and

Ferreira (2006) and Proposition 1 and Theorem 2 in Daouia et al. (2020) that, after straightforward calculations,

$$\sqrt{k_n}\,\overline{r}(1-k_n/n) = \sqrt{k_n}\left[\left(1-\frac{\overline{Y}_n}{\overline{\xi}_{1-k_n/n}}\right)\frac{1}{1-2k_n/n}\left(1+\frac{\overline{b}[\widehat{\overline{F}}_n(\overline{\xi}_{1-k_n/n})]^{-\overline{\rho}}}{1-\overline{\gamma}-\overline{\rho}}\right)^{-1}-1\right]$$
$$= -\frac{(\gamma^{-1}-1)^{-\rho}}{\gamma(1-\gamma-\rho)}\lambda_1 - (\gamma^{-1}-1)^{\gamma}\mathbb{E}[Y]\lambda_2 + \mathrm{o}_{\mathbb{P}}(1). \tag{23}$$

Combine Equations (21), (22), (23) and the delta-method to conclude the proof. □

Proof of Theorem 2. We prove the result when the intermediate and extrapolated expectile estimators featuring in the estimators $\overline{B}_{j,n}$ of the bias terms are the LAWS versions; the proof for the indirect estimators is identical. Clearly

$$\log\left(\frac{\widehat{\xi}^{\star,\mathrm{RB}}_{\beta_n}}{\xi_{\beta_n}}\right)$$
$$= (\overline{\gamma}-\gamma)\log\left(\frac{k_n}{n(1-\beta_n)}\right) + \log\left(\frac{\widehat{\xi}_{1-k_n/n}}{\xi_{1-k_n/n}}\right) - \log\left(\left[\frac{n(1-\beta_n)}{k_n}\right]^{\gamma}\frac{\xi_{\beta_n}}{\xi_{1-k_n/n}}\right)$$
$$+\log(1+\overline{B}_{1,n})+\log(1+\overline{B}_{2,n})+\log(1+\overline{B}_{3,n})$$
$$= (\overline{\gamma}-\gamma)\log\left(\frac{k_n}{n(1-\beta_n)}\right) + \log\left(\frac{\widehat{\xi}_{1-k_n/n}}{\xi_{1-k_n/n}}\right) + \log\left(\frac{1+\overline{B}_{1,n}}{1+B_{1,n}}\right) + \log\left(\frac{1+\overline{B}_{2,n}}{1+B_{2,n}}\right) + \log\left(\frac{1+\overline{B}_{3,n}}{1+B_{3,n}}\right).$$

Similarly,

$$\log\left(\frac{\widetilde{\xi}^{\star,\mathrm{RB}}_{\beta_n}}{\xi_{\beta_n}}\right) = (\overline{\gamma}-\gamma)\log\left(\frac{k_n}{n(1-\beta_n)}\right) + \log\left(\frac{(\overline{\gamma}^{-1}-1)^{-\overline{\gamma}}}{(\gamma^{-1}-1)^{-\gamma}}\right) + \log\left(\frac{Y_{n-k_n,n}}{q_{1-k_n/n}}\right)$$
$$+\log\left(\frac{1+\overline{B}_{1,n}}{1+B_{1,n}}\right)+\log\left(\frac{1+\overline{B}_{3,n}}{1+B_{3,n}}\right).$$

It follows from our assumptions on $\overline{\gamma}$, $\overline{b}$ and $\overline{\rho}$, combined with (for the control of $\overline{B}_{2,n}$ and $\overline{B}_{3,n}$) the law of large numbers and (for the control of $\overline{B}_{2,n}$ specifically in the correction of the LAWS estimator) the Gaussian approximation of the tail empirical distribution function in Theorem 5.1.4 p.161 in de Haan and Ferreira (2006) and Proposition 1 and Theorem 2 in Daouia et al. (2020) that, after straightforward calculations,

$$\log\left(\frac{1+\overline{B}_{1,n}}{1+B_{1,n}}\right)+\log\left(\frac{1+\overline{B}_{2,n}}{1+B_{2,n}}\right)+\log\left(\frac{1+\overline{B}_{3,n}}{1+B_{3,n}}\right) = \mathrm{o}_{\mathbb{P}}\left(\frac{\log(k_n/[n(1-\beta_n)])}{\sqrt{k_n}}\right).$$

The two desired convergences follow immediately (for the convergence of $\widehat{\xi}_{1-k_n/n}$ at the rate $\sqrt{k_n}$, use Theorem 1 in Daouia et al. (2020)). □

C Simulation study results

We give here all supporting tables and figures regarding our simulation study in Section 4.

$\gamma = 0.1$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	-2.05×10^{-3}	-8.98×10^{-4}	1.26×10^{-1}	-2.88×10^{-2}	-1.32×10^{-2}	-2.90×10^{-2}	-1.29×10^{-2}
Variance	3.92×10^{-5}	1.49×10^{-4}	1.10×10^{-3}	3.99×10^{-4}	6.20×10^{-4}	3.10×10^{-4}	6.80×10^{-4}
MSE	4.34×10^{-5}	1.50×10^{-4}	1.70×10^{-2}	1.23×10^{-3}	7.95×10^{-4}	1.15×10^{-3}	8.45×10^{-4}
Average k_n^*	436	49	405	405	49	405	49
$\gamma = 0.2$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	-4.10×10^{-3}	-2.60×10^{-3}	2.29×10^{-1}	-4.32×10^{-2}	-2.64×10^{-2}	-4.37×10^{-2}	-2.48×10^{-2}
Variance	1.57×10^{-4}	3.75×10^{-4}	5.31×10^{-3}	1.83×10^{-3}	3.06×10^{-3}	1.72×10^{-3}	3.18×10^{-3}
MSE	1.74×10^{-4}	3.82×10^{-4}	5.76×10^{-2}	3.70×10^{-3}	3.76×10^{-3}	3.62×10^{-3}	3.80×10^{-3}
Average k_n^*	436	113	405	405	113	405	113
$\gamma = 0.3$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	-6.15×10^{-3}	-4.25×10^{-3}	2.97×10^{-1}	-4.83×10^{-2}	-3.17×10^{-2}	-4.99×10^{-2}	-2.90×10^{-2}
Variance	3.53×10^{-4}	6.97×10^{-4}	1.38×10^{-2}	5.48×10^{-3}	9.34×10^{-3}	5.18×10^{-3}	1.06×10^{-2}
MSE	3.91×10^{-4}	7.14×10^{-4}	1.02×10^{-1}	7.81×10^{-3}	1.03×10^{-2}	7.67×10^{-3}	1.15×10^{-2}
Average k_n^*	436	203	405	405	203	405	203
$\gamma = 0.4$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	-8.20×10^{-3}	-6.16×10^{-3}	3.20×10^{-1}	-4.91×10^{-2}	-2.87×10^{-2}	-5.21×10^{-2}	-2.22×10^{-2}
Variance	6.28×10^{-4}	1.19×10^{-3}	2.71×10^{-2}	1.37×10^{-2}	2.72×10^{-2}	1.22×10^{-2}	3.42×10^{-2}
MSE	6.94×10^{-4}	1.22×10^{-3}	1.30×10^{-1}	1.61×10^{-2}	2.80×10^{-2}	1.49×10^{-2}	3.47×10^{-2}
Average k_n^*	436	333	405	405	333	405	333

Table C.1: Bias, variance and MSE of the tail index and extreme expectile estimators, in the case of the Burr distribution with $\rho = -5$, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$ and (for expectiles) at a level $\beta_n = 1 - 5/n = 0.995$, and $\gamma \in \{0.1, 0.2, 0.3, 0.4\}$. From left to right: (1) bias-reduced Hill estimator $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$, (2) bias-reduced asymptotic proportionality tail index estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$, (3) standard extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (4) bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (5) bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step, (6) bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (7) bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step. For expectile estimates, relative biases, variances and MSEs are reported. Blue highlighting indicates the best bias/variance/MSE among the two tail index estimates, while violet highlighting indicates the best bias/variance/MSE among the five extreme expectile estimates.

$\gamma = 0.1$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	-2.18×10^{-3}	-1.23×10^{-3}	7.62×10^{-2}	-8.05×10^{-3}	-3.50×10^{-4}	-8.54×10^{-3}	-3.12×10^{-3}
Variance	1.14×10^{-4}	3.36×10^{-4}	2.10×10^{-3}	6.12×10^{-4}	8.49×10^{-4}	6.05×10^{-4}	1.09×10^{-3}
MSE	1.19×10^{-4}	3.37×10^{-4}	7.90×10^{-3}	6.77×10^{-4}	8.48×10^{-4}	6.78×10^{-4}	1.09×10^{-3}
Average k_n^*	108	12	108	108	12	108	12
$\gamma = 0.2$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	-4.37×10^{-3}	-3.37×10^{-3}	1.40×10^{-1}	-1.22×10^{-2}	-5.05×10^{-3}	-1.31×10^{-2}	-6.40×10^{-3}
Variance	4.56×10^{-4}	8.46×10^{-4}	1.04×10^{-2}	3.72×10^{-3}	4.83×10^{-3}	3.76×10^{-3}	5.15×10^{-3}
MSE	4.74×10^{-4}	8.56×10^{-4}	3.00×10^{-2}	3.87×10^{-3}	4.85×10^{-3}	3.93×10^{-3}	5.19×10^{-3}
Average k_n^*	108	30	108	108	30	108	30
$\gamma = 0.3$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	-6.55×10^{-3}	-5.83×10^{-3}	1.90×10^{-1}	-1.15×10^{-2}	-5.81×10^{-3}	-1.22×10^{-2}	-5.00×10^{-3}
Variance	1.02×10^{-3}	1.55×10^{-3}	2.91×10^{-2}	1.25×10^{-2}	1.58×10^{-2}	1.27×10^{-2}	1.63×10^{-2}
MSE	1.07×10^{-3}	1.58×10^{-3}	6.51×10^{-2}	1.26×10^{-2}	1.58×10^{-2}	1.28×10^{-2}	1.63×10^{-3}
Average k_n^*	108	57	108	108	57	108	57
$\gamma = 0.4$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	-8.74×10^{-3}	-8.13×10^{-2}	2.25×10^{-1}	-6.84×10^{-3}	-1.29×10^{-3}	-5.64×10^{-3}	5.68×10^{-3}
Variance	1.82×10^{-3}	2.32×10^{-3}	6.58×10^{-2}	3.40×10^{-2}	4.27×10^{-2}	3.38×10^{-2}	4.70×10^{-2}
MSE	1.90×10^{-3}	2.38×10^{-3}	1.16×10^{-1}	3.40×10^{-2}	4.26×10^{-2}	3.37×10^{-2}	4.69×10^{-2}
Average k_n^*	108	115	108	108	115	108	115

Table C.2: Bias, variance and MSE of the tail index and extreme expectile estimators, in the case of the Burr distribution with $\rho = -1$, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$ and (for expectiles) at a level $\beta_n = 1 - 5/n = 0.995$, and $\gamma \in \{0.1, 0.2, 0.3, 0.4\}$. From left to right: (1) bias-reduced Hill estimator $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$, (2) bias-reduced asymptotic proportionality tail index estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$, (3) standard extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (4) bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (5) bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step, (6) bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (7) bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step. For expectile estimates, relative biases, variances and MSEs are reported. Blue highlighting indicates the best bias/variance/MSE among the two tail index estimates, while violet highlighting indicates the best bias/variance/MSE among the five extreme expectile estimates.

$\gamma = 0.1$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	1.43×10^{-2}	9.50×10^{-3}	7.46×10^{-2}	3.43×10^{-3}	2.76×10^{-3}	3.99×10^{-3}	1.66×10^{-3}
Variance	1.41×10^{-4}	3.48×10^{-4}	2.21×10^{-3}	9.58×10^{-4}	1.12×10^{-3}	9.09×10^{-4}	1.47×10^{-3}
MSE	3.46×10^{-4}	4.38×10^{-4}	7.77×10^{-3}	9.69×10^{-4}	1.13×10^{-3}	9.24×10^{-4}	1.47×10^{-3}
Average k_n^*	84	11	84	84	11	84	11
$\gamma = 0.2$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	2.87×10^{-2}	2.19×10^{-2}	1.51×10^{-1}	1.84×10^{-2}	9.92×10^{-3}	2.01×10^{-2}	1.39×10^{-2}
Variance	5.64×10^{-4}	8.94×10^{-4}	1.20×10^{-2}	5.88×10^{-3}	6.83×10^{-3}	6.02×10^{-3}	7.32×10^{-3}
MSE	1.39×10^{-3}	1.37×10^{-3}	3.48×10^{-2}	6.22×10^{-3}	6.92×10^{-3}	6.42×10^{-3}	7.50×10^{-3}
Average k_n^*	84	28	84	84	28	84	28
$\gamma = 0.3$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	4.30×10^{-2}	3.60×10^{-2}	2.32×10^{-1}	4.80×10^{-2}	3.45×10^{-2}	5.77×10^{-2}	4.04×10^{-2}
Variance	1.27×10^{-3}	1.53×10^{-3}	3.85×10^{-2}	2.08×10^{-2}	2.29×10^{-2}	2.27×10^{-2}	2.29×10^{-2}
MSE	3.12×10^{-3}	2.83×10^{-3}	9.24×10^{-2}	2.30×10^{-2}	2.40×10^{-2}	2.60×10^{-2}	2.45×10^{-2}
Average k_n^*	84	61	84	84	61	84	61
$\gamma = 0.4$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	5.73×10^{-2}	5.23×10^{-2}	3.24×10^{-1}	9.31×10^{-2}	8.63×10^{-2}	1.29×10^{-1}	5.62×10^{-2}
Variance	2.26×10^{-3}	2.57×10^{-3}	1.05×10^{-1}	6.16×10^{-2}	7.63×10^{-2}	7.39×10^{-2}	9.28×10^{-2}
MSE	5.54×10^{-3}	5.30×10^{-3}	2.10×10^{-1}	7.02×10^{-2}	8.36×10^{-2}	9.06×10^{-2}	9.59×10^{-2}
Average k_n^*	84	179	84	84	179	84	179

Table C.3: Bias, variance and MSE of the tail index and extreme expectile estimators, in the case of the Burr distribution with $\rho = -0.5$, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$ and (for expectiles) at a level $\beta_n = 1 - 5/n = 0.995$, and $\gamma \in \{0.1, 0.2, 0.3, 0.4\}$. From left to right: (1) bias-reduced Hill estimator $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$, (2) bias-reduced asymptotic proportionality tail index estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$, (3) standard extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (4) bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (5) bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step, (6) bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (7) bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step. For expectile estimates, relative biases, variances and MSEs are reported. Blue highlighting indicates the best bias/variance/MSE among the two tail index estimates, while violet highlighting indicates the best bias/variance/MSE among the five extreme expectile estimates.

$\gamma = 0.1$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	2.22×10^{-1}	2.03×10^{-1}	2.72×10^{-1}	1.25×10^{-1}	8.27×10^{-2}	1.49×10^{-1}	1.07×10^{-1}
Variance	9.95×10^{-4}	9.07×10^{-4}	2.48×10^{-2}	1.37×10^{-2}	1.18×10^{-2}	1.70×10^{-2}	1.20×10^{-2}
MSE	5.02×10^{-2}	4.23×10^{-2}	9.86×10^{-2}	2.93×10^{-2}	1.87×10^{-2}	3.92×10^{-2}	2.34×10^{-2}
Average k_n^*	78	49	78	78	49	78	49
$\gamma = 0.2$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	1.77×10^{-1}	1.60×10^{-1}	3.26×10^{-1}	1.42×10^{-1}	1.01×10^{-1}	1.79×10^{-1}	1.05×10^{-1}
Variance	1.41×10^{-3}	1.26×10^{-3}	4.45×10^{-2}	2.46×10^{-2}	2.28×10^{-2}	3.25×10^{-2}	1.93×10^{-2}
MSE	3.27×10^{-2}	2.69×10^{-2}	1.51×10^{-1}	4.48×10^{-2}	3.30×10^{-2}	6.44×10^{-2}	3.03×10^{-2}
Average k_n^*	80	75	80	80	75	80	75
$\gamma = 0.3$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	1.37×10^{-1}	1.28×10^{-1}	3.74×10^{-1}	1.53×10^{-1}	1.36×10^{-1}	2.08×10^{-1}	1.00×10^{-1}
Variance	1.94×10^{-3}	2.21×10^{-3}	7.83×10^{-2}	4.41×10^{-2}	5.14×10^{-2}	6.05×10^{-2}	3.81×10^{-2}
MSE	2.08×10^{-2}	1.85×10^{-2}	2.18×10^{-1}	6.76×10^{-2}	6.98×10^{-2}	1.04×10^{-1}	4.81×10^{-2}
Average k_n^*	81	141	81	81	141	81	141
$\gamma = 0.4$	$\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$	$\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$	$\widehat{\xi}_{\beta_n}^{\star}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widehat{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{H,RB}})$	$\widetilde{\xi}_{\beta_n}^{\star,\mathrm{RB}}\ (\widehat{\gamma}_{k_n}^{\mathrm{E,RB}})$
Bias	1.02×10^{-1}	1.02×10^{-1}	4.11×10^{-1}	1.58×10^{-1}	1.66×10^{-1}	2.37×10^{-1}	8.02×10^{-2}
Variance	2.66×10^{-3}	2.36×10^{-3}	1.46×10^{-1}	8.59×10^{-2}	8.75×10^{-2}	1.27×10^{-1}	9.09×10^{-2}
MSE	1.31×10^{-2}	1.28×10^{-2}	3.14×10^{-1}	1.11×10^{-1}	1.15×10^{-1}	1.83×10^{-1}	9.72×10^{-2}
Average k_n^*	82	219	82	82	219	82	219

Table C.4: Bias, variance and MSE of the tail index and extreme expectile estimators, in the case of the Generalised Pareto Distribution, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$ and (for expectiles) at a level $\beta_n = 1 - 5/n = 0.995$, and $\gamma \in \{0.1, 0.2, 0.3, 0.4\}$. From left to right: (1) bias-reduced Hill estimator $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$, (2) bias-reduced asymptotic proportionality tail index estimator $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$, (3) standard extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (4) bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (5) bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step, (6) bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step, (7) bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step. For expectile estimates, relative biases, variances and MSEs are reported. Blue highlighting indicates the best bias/variance/MSE among the two tail index estimates, while violet highlighting indicates the best bias/variance/MSE among the five extreme expectile estimates.

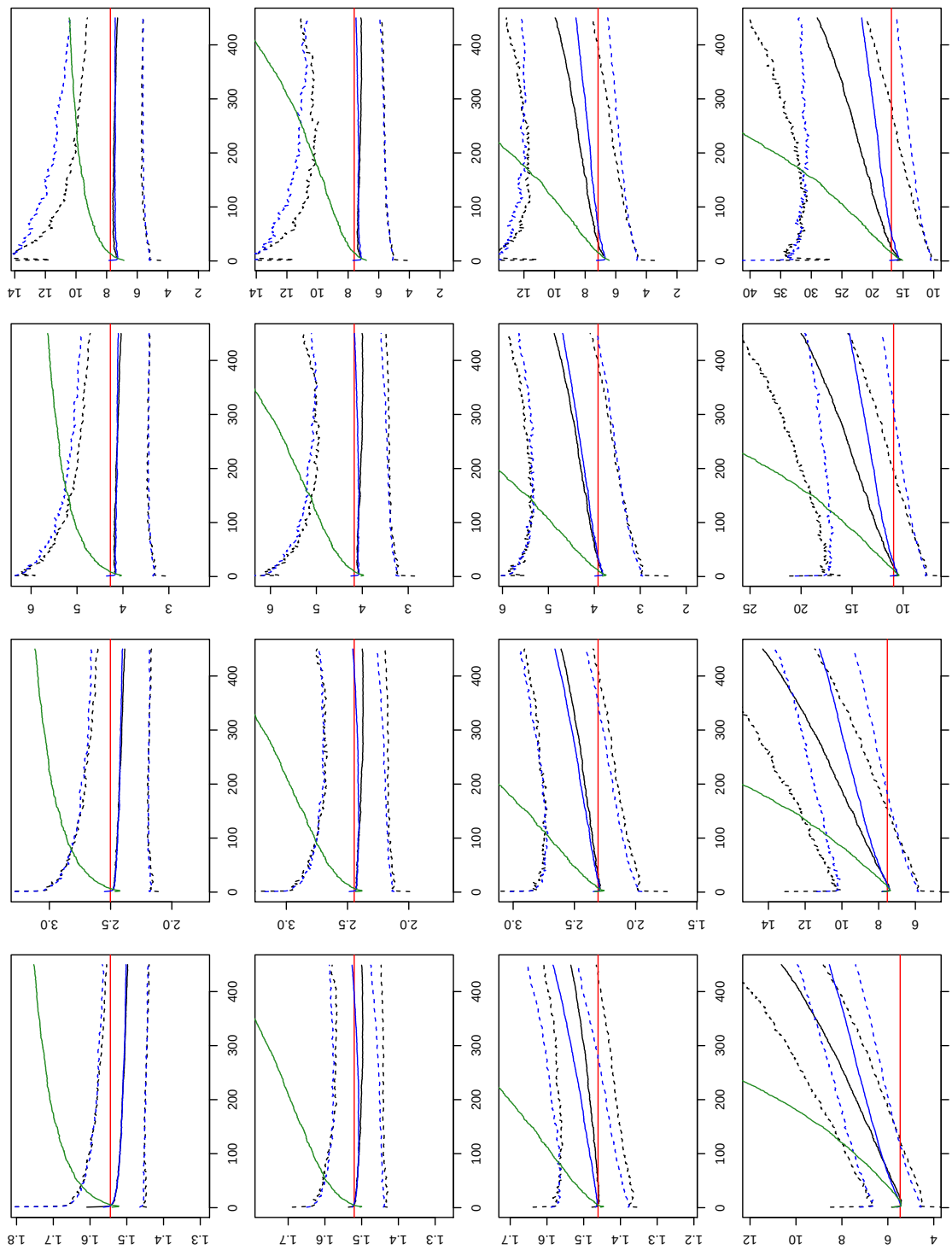

Figure C.1: Median extreme expectile estimates, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$ at a level $\beta_n = 1 - 5/n = 0.995$, and for $k_n \in \{2, 3, \ldots, 450\}$. The solid lines represent the standard extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step (green), the bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step (black) and the bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step (blue), with the red line representing the true value of the expectile to be estimated. The dotted lines represent empirical 95% confidence intervals based on the bias-reduced estimates with the same colour code. From top to bottom: Burr distribution with $\rho = -5$, Burr distribution with $\rho = -1$, Burr distribution with $\rho = -0.5$, and Generalised Pareto Distribution; from left to right, $\gamma = 0.1, 0.2, 0.3, 0.4$.

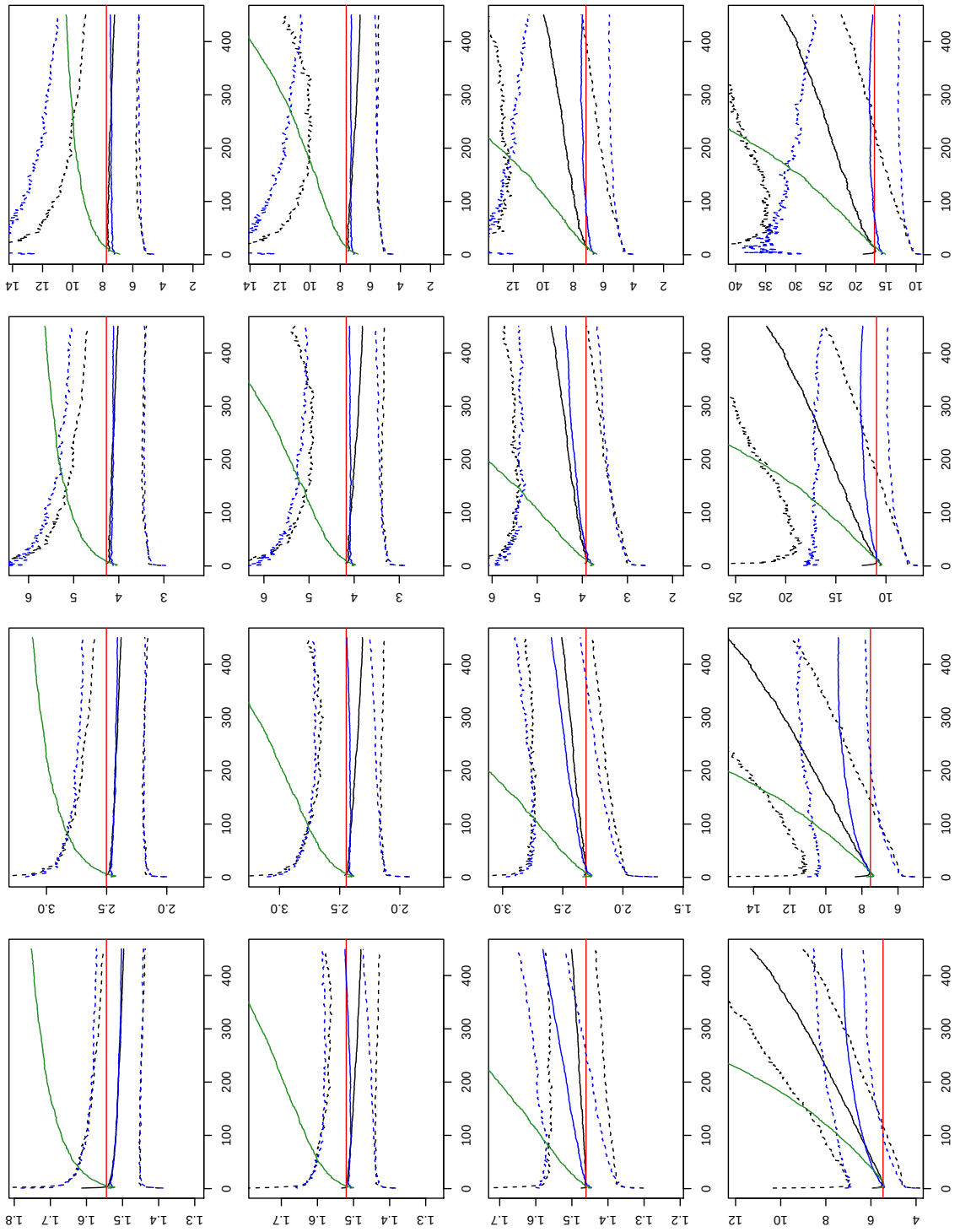

Figure C.2: Median extreme expectile estimates, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$ at a level $\beta_n = 1 - 5/n = 0.995$, and for $k_n \in \{2, 3, \ldots, 450\}$. The solid lines represent the standard extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step (green), the bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step (black) and the bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step (blue), with the red line representing the true value of the expectile to be estimated. The dotted lines represent empirical 95% confidence intervals based on the bias-reduced estimates with the same colour code. From top to bottom: Burr distribution with $\rho = -5$, Burr distribution with $\rho = -1$, Burr distribution with $\rho = -0.5$, and Generalised Pareto Distribution; from left to right, $\gamma = 0.1, 0.2, 0.3, 0.4$.

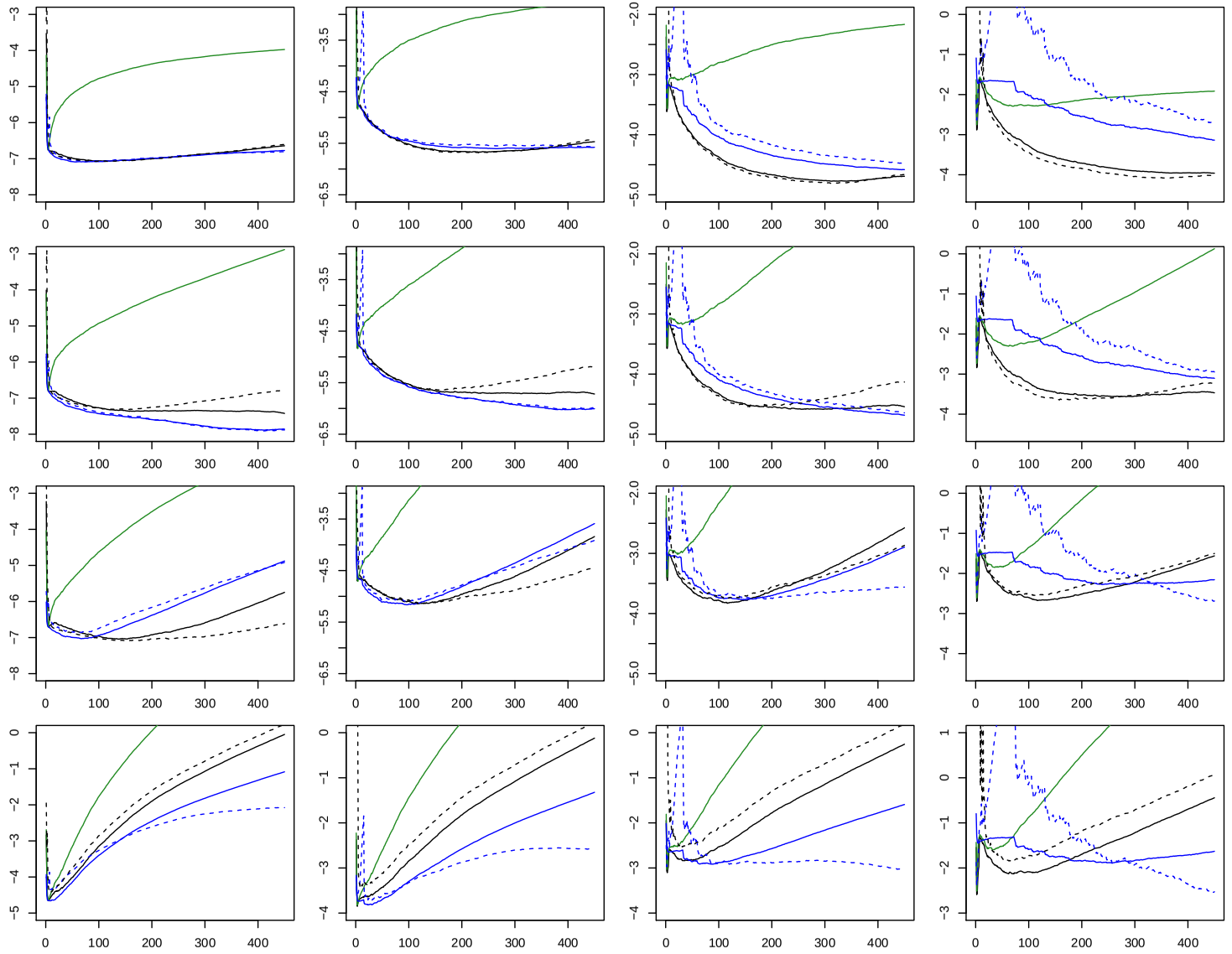

Figure C.3: Log relative mean-squared errors of extreme expectile estimators, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$ at a level $\beta_n = 1 - 5/n = 0.995$, and for $k_n \in \{2, 3, \ldots, 450\}$. The solid lines represent the standard extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step (green), the bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step (black) and the bias-reduced extrapolated LAWS expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step (blue). The dotted lines represent the bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{H,RB}}$ at the extrapolation step (black) and the bias-reduced extrapolated indirect expectile estimator using $\widehat{\gamma}_{k_n}^{\mathrm{E,RB}}$ at the extrapolation step (blue). From top to bottom: Burr distribution with $\rho = -5$, Burr distribution with $\rho = -1$, Burr distribution with $\rho = -0.5$, and Generalised Pareto Distribution; from left to right, $\gamma = 0.1, 0.2, 0.3, 0.4$.

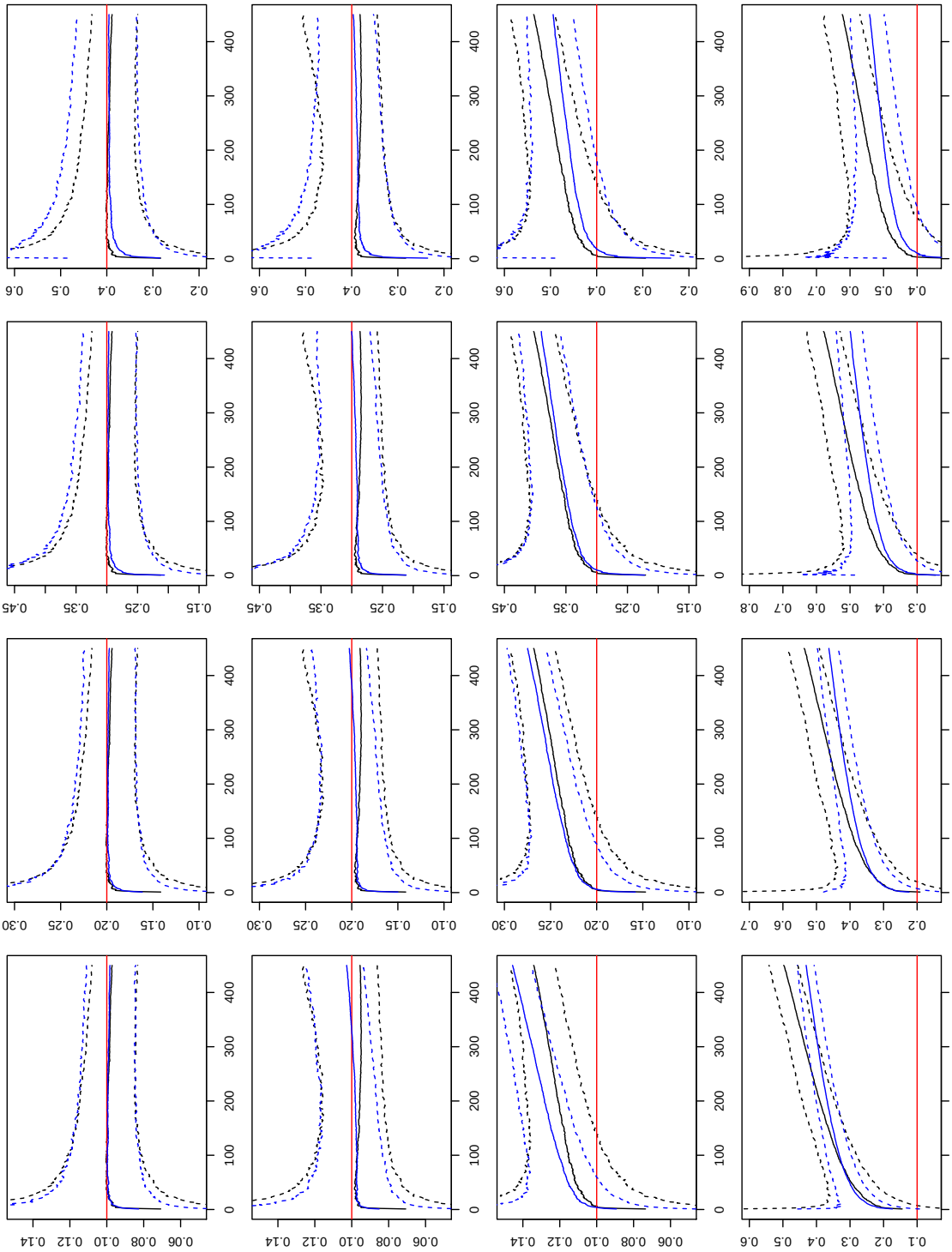

Figure C.4: Median tail index estimates, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$, and for $k_n \in \{2, 3, \ldots, 450\}$. The solid lines represent the bias-reduced Hill estimates (black) and bias-reduced asymptotic proportionality estimates (blue), with the red line representing the true value. The dotted lines represent empirical 95% confidence intervals with the same colour code. From top to bottom: Burr distribution with $\rho = -5$, Burr distribution with $\rho = -1$, Burr distribution with $\rho = -0.5$, and Generalised Pareto Distribution; from left to right, $\gamma = 0.1, 0.2, 0.3, 0.4$.

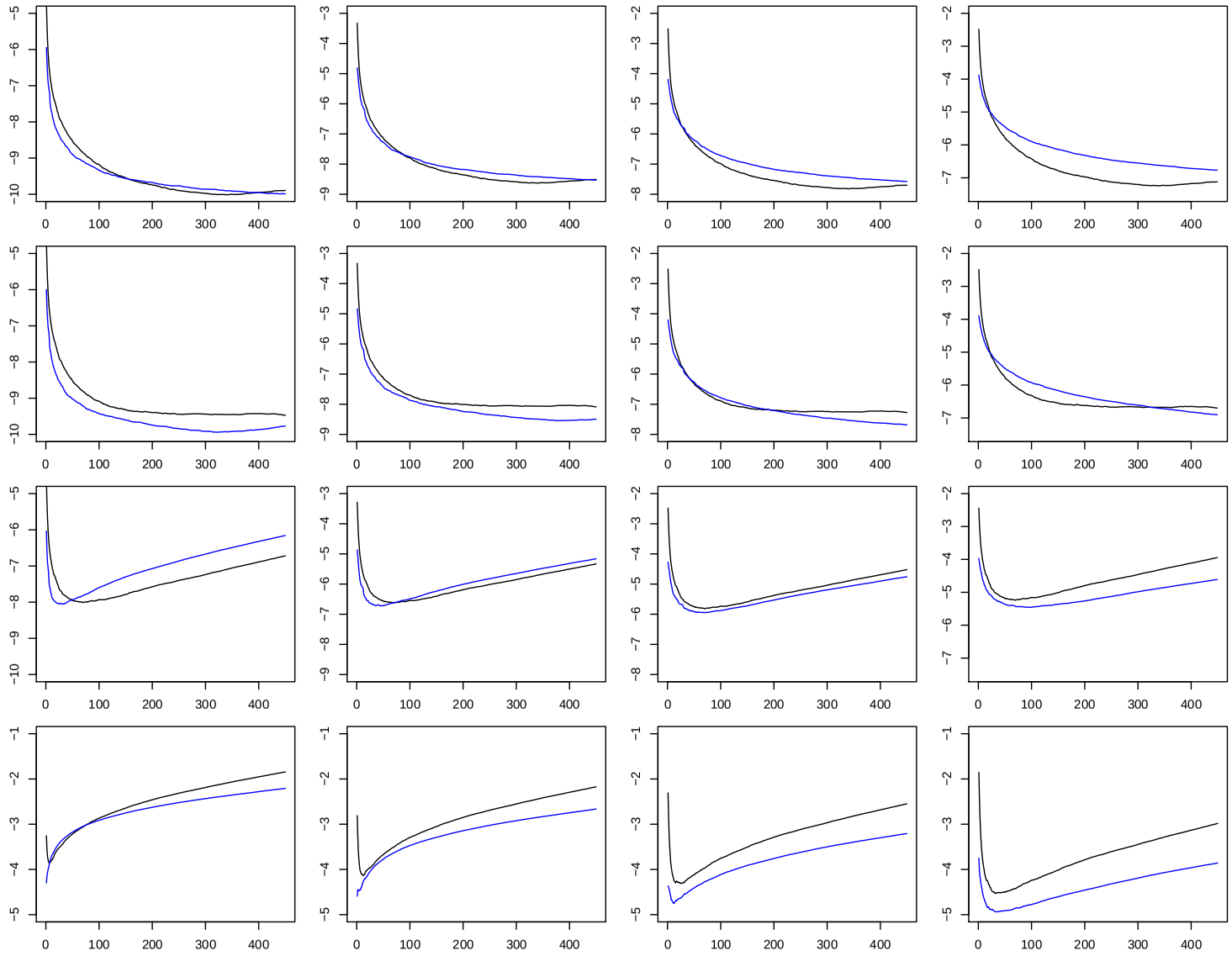

Figure C.5: Log mean-squared errors of tail index estimators, for $N = 1{,}000$ replications of a sample of size $n = 1{,}000$, and for $k_n \in \{2, 3, \ldots, 450\}$. The solid lines represent the bias-reduced Hill estimator (black) and bias-reduced asymptotic proportionality estimator (blue). From top to bottom: Burr distribution with $\rho = -5$, Burr distribution with $\rho = -1$, Burr distribution with $\rho = -0.5$, and Generalised Pareto Distribution; from left to right, $\gamma = 0.1, 0.2, 0.3, 0.4$.