*Submitted to Bernoulli*

# Joint inference on extreme expectiles for multivariate heavy-tailed distributions

SIMONE A. PADOAN[1], and GILLES STUPFLER[2,*]

[1] *Department of Decision Sciences, Bocconi University, via Roentgen 1, 20136 Milano, Italy*
*E-mail:* simone.padoan@unibocconi.it
[2] *Univ Rennes, Ensai, CNRS, CREST - UMR 9194, Rennes, France*
*E-mail:* [*]gilles.stupfler@ensai.fr

Expectiles induce a law-invariant, coherent and elicitable risk measure that has received substantial attention in actuarial and financial risk management contexts. A number of recent papers have focused on the behaviour and estimation of extreme expectile-based risk measures and their potential for risk assessment was highlighted in financial and actuarial real data applications. Joint inference of several extreme expectiles has however been left untouched; in fact, even the inference about a marginal extreme expectile turns out to be a difficult problem in finite samples, even though an accurate idea of estimation uncertainty is crucial for the construction of confidence intervals in applications to risk management. We investigate the joint estimation of extreme marginal expectiles of a random vector with heavy-tailed marginal distributions, in a general extremal dependence model. We use these results to derive corrected confidence regions for extreme expectiles, as well as a test for the equality of tail expectiles. The methods are showcased in a finite-sample simulation study and on real financial data.

*Keywords:* Expectiles; Extremal dependence; Heavy tails; Joint convergence; Joint inference; Tail copula; Testing

## 1. Introduction and background

Expectiles, introduced by [28], induce risk measures which have recently gained substantial traction in the risk management context. Expectiles of an integrable random variable $X$ are obtained as minimisers of asymmetrically squared deviations in the following sense:

$$\xi_\tau = \underset{\theta \in \mathbb{R}}{\arg\min}\, \mathbb{E}(\eta_\tau(X-\theta) - \eta_\tau(X)), \tag{1}$$

where $\eta_\tau(u) = |\tau - \mathbb{1}\{u \le 0\}|u^2$ is the so-called expectile check function and $\mathbb{1}\{\cdot\}$ the indicator function. Expectiles can be seen as $L^2-$analogues of quantiles, which can be obtained by minimising asymmetrically weighted mean absolute deviations [26], *i.e.* by considering the quantile check function $\rho_\tau(u) = |\tau - \mathbb{1}\{u \le 0\}||u|$ instead of $\eta_\tau$. By construction, the $\tau$th expectile depends on tail realisations of $X$ as well as their probability. The advantages of the expectile include that it is the only risk measure, apart from the simple expectation, that defines a law-invariant, coherent [1] and elicitable [17] risk measure, see [3] and [36]. Being elicitable, expectiles benefit from the existence of a natural backtesting methodology. Quantiles are elicitable too, but are often criticised for not being a coherent risk measure, and for missing out on important information about the tail of the underlying distribution since they only depend on the frequency of tail events. Meanwhile, the popular quantile-based Expected Shortfall is coherent, takes into account the actual values of the risk variable on the tail event, but is not elicitable. Further investigations carried out by [2, 15], among others, suggest that expectiles define perfectly sensible alternatives to the quantile and Expected Shortfall.

Expectile estimation has regained interest in a large range of models, see for example [31] for the estimation of central expectiles of fixed order $\tau$. Probabilistic and statistical aspects of extreme expectiles, with $\tau \uparrow 1$, have been examined by [2, 3] and [7, 8, 9] respectively. This last strand of work on estimation of extreme expectiles finds nice applications in the assessment of extreme risk, whether at the company level or at the systemic level. However, these estimation results are limited to the consideration of extreme expectiles of a single sample of data, and hence to a single asset or risk variable; this is a substantial restriction in actuarial and financial applications, where practitioners are interested in evaluating the asymptotic dependence existing within several risk variables, stock prices or stock indices, and in carrying out precise joint inference about the extremes of these risk variables. Such questions are for instance considered in [32] regarding the detection of tail asymmetries, [27] for the construction of diversified financial portfolios, and in [23] as a way to directly compare risk measures between assets. Besides, an inspection of the Gaussian QQ-plots in Appendix A.2 of [7] shows that, despite the fact that the Gaussian distribution will in many cases be a reasonable model for the uncertainty of extreme expectile estimators, the sample variance of the estimators can be a long way off the variance obtained *via* a naive use of the theoretical Gaussian approximation. This has made using expectiles very difficult in applications to risk management, where an accurate idea of estimation uncertainty is crucial for the construction of confidence intervals, based on currently available techniques.

This paper makes a step towards filling that gap in the following ways. In a general framework of multivariate distributions with marginal heavy tails and extremal pairwise dependence between margins, and given independent and identically distributed (i.i.d.) data, we start by rigorously investigating the joint asymptotic normality of extreme expectile estimators of the margins. Our assumptions are mild: in particular, we avoid assuming a particular form of multivariate extreme dependence structure. These results are then used to tackle the question of joint inference about tail expectiles from two distinct angles. First, we exploit the joint Gaussian asymptotics of tail expectile estimators to construct asymptotic joint confidence regions for tail expectiles. This is done by, on the one hand, designing specific finite-sample corrections for the standard plug-in asymptotic variance estimators of each expectile estimator to obtain accurate representations of marginal uncertainty. On the other hand, we construct an appropriate nonparametric estimator of the tail dependence between two such estimators pertaining to different marginals. This results in an estimate of the covariance matrix of our set of expectile estimators, used to build Gaussian-based confidence regions for the vector of expectiles of interest. Our procedure is computationally very fast and avoids resorting to either bootstrapping, whose calibration and justification with a heavy-tailed underlying distribution may be very difficult, or an Edgeworth expansion, which for expectiles (that extend the mean) would intuitively require strong moment assumptions that we aim to avoid in our heavy-tailed context. Second, we tackle the problem of testing whether tail expectiles across marginals are equal. We do so by adapting the classical likelihood ratio test of equal means in a Gaussian random vector. The deviance test statistic prominently features our covariance matrix estimators that will be used to construct accurate confidence regions. This testing procedure should be viewed as an addition to existing multivariate tools, such as the procedures in [22] based on extreme quantiles, whose aim is to test whether different distributions have the same extreme value behaviour. This shall be illustrated in our financial data application on currency exchange rates.

The outline of the paper is the following. Section 2 explains our statistical context and contains the main theoretical results of the paper on joint intermediate and extreme expectile estimation. We then develop our joint inference methods about tail expectiles in Section 3. The finite-sample performance of the methods is examined on simulated data sets in Section 4 and on financial exchange rates data in Section 5. We discuss our findings and perspectives for future work in Section 6. The methods and data considered in this article have been incorporated into the freely available `R` package `ExtremeRisks`. The Supplementary Material document accompanying this article contains all mathematical proofs, further details on our construction of confidence regions, and further finite-sample results.

## 2. Joint estimation of multiple extreme expectiles

Let $(\boldsymbol{X}_i, 1 \le i \le n)$, with $\boldsymbol{X}_i = (X_{i,j}, 1 \le j \le d)$, be i.i.d. copies of a $d$-dimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_d)$, with marginal distributions $F_j$, associated survival functions $\overline{F}_j = 1 - F_j$, and tail quantile functions $U_j(s) = \inf\{x \in \mathbb{R} \,|\, F_j(x) \ge 1 - s^{-1}\}$, for $s > 1$. The realisations of $X_j$ may for example be seen as the negatives of generic financial positions, so that large positive values of $X_j$ represent extreme losses, or as losses incurred by an insurance company in distinct lines of business.

We focus on the joint estimation of extreme expectiles of $X_1, \ldots, X_d$. We work with heavy-tailed distributions, representing the tail structure of many financial and actuarial data examples fairly well, see *e.g.* p.9 of [16]. Mathematically, we assume

$$\forall j \in \{1, \ldots, d\},\ \forall x > 0,\ \lim_{s\to\infty} \frac{\overline{F}_j(sx)}{\overline{F}_j(s)} = x^{-1/\gamma_j} \text{ or equivalently } \lim_{s\to\infty} \frac{U_j(sx)}{U_j(s)} = x^{\gamma_j}.$$

The tail indices $\gamma_j > 0$ specify marginal tail heaviness. With condition $\mathbb{E}|\min(X_j, 0)| < \infty$, the assumption $0 < \gamma_j < 1$ ensures that the first moment of $X_j$ exists and thus expectiles of the $X_j$ are well-defined. These two conditions will be part of our assumptions throughout.

We seek to establish the joint asymptotic distribution of tail expectile estimators of level $\tau$ close to 1. Specifically, according to (1), the expectile for the $j$th marginal distribution $F_j$ is defined as

$$\xi_{\tau,j} = \underset{\theta\in\mathbb{R}}{\arg\min}\, \mathbb{E}(\eta_\tau(X_j - \theta) - \eta_\tau(X_j)), \tag{2}$$

where $\eta_\tau$ is the expectile check function defined below Equation (1). We consider hereafter the problem of the joint inference of $(\xi_{\tau,1}, \ldots, \xi_{\tau,d})$, where the level $\tau$ is such that $\tau = 1 - p$ for a small value of $p = p_n$. Two cases are considered, when $p \gg 1/n$ and $p \approx c/n$ with $c$ a finite constant: these are respectively the *intermediate case*, when nonparametric estimation methods can be used, and the properly *extreme case* when extrapolation methods whose rationale is rooted in the heavy-tailed assumption have to be developed. To carry out joint inference about estimators of extreme expectiles, we model here the extremal dependence structure between any two components of $\boldsymbol{X}$ in the form of a tail copula. This translates into the following general assumption that we shall work with throughout.

***Condition* A.** *For every $1 \le j \le d$, let $F_j$ and $U_j$ be the marginal distribution function and tail quantile function associated to $X_j$. Assume that the $F_j$ are continuous and:*

*(i) $U_j$ is regularly varying with index $\gamma_j$: $U_j(sx)/U_j(s) \to x^{\gamma_j}$ as $s \to \infty$, for any $x > 0$.*

*(ii) For any $(j, \ell)$ with $j \ne \ell$, there is a function $R_{j,\ell}$ on $[0,\infty]^2 \setminus \{(\infty,\infty)\}$ such that*

$$\forall (x_j, x_\ell) \in [0,\infty]^2 \setminus \{(\infty,\infty)\},\ \lim_{s\to\infty} s\,\mathbb{P}\left(\overline{F}_j(X_j) \le \frac{x_j}{s},\ \overline{F}_\ell(X_\ell) \le \frac{x_\ell}{s}\right) = R_{j,\ell}(x_j, x_\ell).$$

Condition A(ii) formalises the existence of a limiting dependence structure in the upper tail of any two components $X_j$ and $X_\ell$, given by the *tail copula* $R_{j,\ell}$ (see [30]). It is a weak assumption, satisfied by any $\boldsymbol{X}$ in the maximum domain of attraction of a multivariate extreme value distribution (see [10]).

### 2.1. At the intermediate level

Let $\tau_n \in (0,1)$ satisfy $\tau_n \to 1$ and $n(1-\tau_n) \to \infty$ as $n \to \infty$. We estimate expectiles of the $X_j$ at level $\tau_n$ using the empirical counterpart of (2), called the *Least Asymmetrically Weighted Squares* (LAWS) estimator, or a semiparametric *Quantile-Based* (QB) estimator built on our heavy-tailed assumption.

**Nonparametric estimator via asymmetric least squares** We first consider estimating the expectile $\xi_{\tau_n,j}$ of the marginal distribution $F_j$ by its empirical (LAWS) estimator

$$\widetilde{\xi}_{\tau_n,j} = \arg\min_{\theta\in\mathbb{R}} \sum_{i=1}^{n} \eta_{\tau_n}(X_{i,j}-\theta).$$

Theorem 2 in [7] shows that the estimator $\widetilde{\xi}_{\tau_n,j}$ is $\sqrt{n(1-\tau_n)}$−asymptotically normal; this result is limited to the marginal estimation of an intermediate expectile. The first main result of this paper provides the joint asymptotic normality of the $\widetilde{\xi}_{\tau_n,j}$, for $1\le j\le d$. The main idea for a proof is to write

$$\sqrt{n(1-\tau_n)}\left(\frac{\widetilde{\xi}_{\tau_n,j}}{\xi_{\tau_n,j}}-1\right) = \arg\min_{u\in\mathbb{R}} \psi_n^{(j)}(u) \tag{3}$$

$$\text{with } \psi_n^{(j)}(u) = \frac{1}{2\xi_{\tau_n,j}^2}\sum_{i=1}^{n}\left[\eta_{\tau_n}\left(X_{i,j}-\xi_{\tau_n,j}-\frac{u\,\xi_{\tau_n,j}}{\sqrt{n(1-\tau_n)}}\right)-\eta_{\tau_n}(X_{i,j}-\xi_{\tau_n,j})\right].$$

The joint convergence of the random functions $u\mapsto\psi_n^{(j)}(u)$ (for $1\le j\le d$) then guarantees the joint convergence of their minimisers $\widetilde{\xi}_{\tau_n,j}$, for example by Theorem 5 in [25]. Denote by $\xrightarrow{\mathrm{d}}$ and $\xrightarrow{\mathbb{P}}$ weak convergence and convergence in probability of sequences of random variables, respectively.

**Theorem 2.1.** *Assume that Condition A is satisfied. Assume further that there is $\delta>0$ such that $\mathbb{E}|\min(X_j,0)|^{2+\delta}<\infty$ and that $0<\gamma_j<1/2$ for any $1\le j\le d$. Let $\tau_n\uparrow 1$ be such that $n(1-\tau_n)\to\infty$ as $n\to\infty$. Then we have*

$$\sqrt{n(1-\tau_n)}\left(\frac{\widetilde{\xi}_{\tau_n,j}}{\xi_{\tau_n,j}}-1\right)_{1\le j\le d} \xrightarrow{\mathrm{d}} \mathcal{N}_d\left(\mathbf{0}_d, \boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})\right).$$

*The covariance matrix $\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ has entries*

$$\boldsymbol{V}_{j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R}) = \begin{cases} \dfrac{2\gamma_j^3}{1-2\gamma_j} & (j=\ell), \\ \gamma_j\gamma_\ell \displaystyle\iint_{[1,\infty)^2} R_{j,\ell}\left((\gamma_j^{-1}-1)x_j^{-1/\gamma_j},(\gamma_\ell^{-1}-1)x_\ell^{-1/\gamma_\ell}\right)\mathrm{d}x_j\,\mathrm{d}x_\ell & (j<\ell). \end{cases}$$

Theorem 2.1 is the extension of Theorem 2 in [7] needed to carry out multivariate inference about intermediate expectiles. Like in the latter, it is assumed that a finite $(2+\delta)$−moment of $X_j$ exists, which is natural because the proof requires showing the asymptotic normality of a tail average of the $X_{i,j}$ $(1\le i\le n)$. Theorem 2.1 is derived under the additional Condition A(ii) which makes it possible to evaluate the asymptotic dependence between the $\widetilde{\xi}_{\tau_n,j}$ through the tail copulae $R_{j,\ell}$. This will be a crucial element for the construction of joint confidence regions for extreme expectiles in Section 3.

**Semiparametric estimator via a quantile-based procedure** An alternative estimator is provided by the asymptotic proportionality relationship between expectile and quantile first noted by [3]:

$$\lim_{\tau\uparrow 1}\frac{\xi_{\tau,j}}{q_{\tau,j}} = (\gamma_j^{-1}-1)^{-\gamma_j}, \tag{4}$$

where $q_{\tau,j}$ is the $\tau$th quantile of the $j$th marginal. This suggests the class of QB estimators

$$\widehat{\xi}_{\tau_n,j} = (\widehat{\gamma}_{\tau_n,j}^{-1} - 1)^{-\widehat{\gamma}_{\tau_n,j}} \widehat{q}_{\tau_n,j}$$

where for each $j \in \{1,\dots,d\}$, $\widehat{q}_{\tau_n,j}$ and $\widehat{\gamma}_{\tau_n,j}$ are consistent estimators of $q_{\tau_n,j}$ and $\gamma_j$. We take throughout $\widehat{q}_{\tau_n,j} = X_{n-\lfloor n(1-\tau_n)\rfloor,n,j}$, where $\lfloor\cdot\rfloor$ is the floor function and $X_{1,n,j} \le \cdots \le X_{n,n,j}$ denote the order statistics of the sample $(X_{1,j},\dots X_{n,j})$. The estimator of $\gamma_j$ will be the Hill estimator introduced in [21], with effective sample size $k = \lfloor n(1-\tau_n)\rfloor$:

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

This estimator is the maximum likelihood estimator in the purely Pareto model. It is minimax optimal and attains the sharp minimax bound in the wide Hall-Welsh class of models of [20], *i.e.* when

$$\overline{F}_j(x) = x^{-1/\gamma_j}\left(a_j + b_j x^{\rho_j/\gamma_j} + \mathrm{o}(x^{\rho_j/\gamma_j})\right) \text{ as } x \to \infty, \tag{5}$$

where $a_j > 0$, $b_j \ne 0$ and $\rho_j < 0$. See Theorems 2.1 and 2.2 in [12].

The asymptotic normality of a single one of the $\widehat{\xi}_{\tau_n,j}$ is investigated in Corollary 2 of [7]. To write the corresponding joint convergence result, we require the following set of second-order conditions.

***Condition* B.** *Assume that Condition A(i) holds and that, for every $1 \le j \le d$,*

$$\forall x > 0,\ \lim_{s\to\infty} \frac{1}{A_j(s)}\left(\frac{U_j(sx)}{U_j(s)} - x^{\gamma_j}\right) = x^{\gamma_j}\frac{x^{\rho_j}-1}{\rho_j},$$

*where $\rho_j \le 0$ and $A_j$ is a measurable function converging to 0 at infinity and having constant sign. Hereafter, $(x^{\rho_j}-1)/\rho_j$ is to be read as $\log(x)$ when $\rho_j = 0$.*

Condition B controls rates of convergences in Condition A(i): since $|A_j|$ is regularly varying with index $\rho_j$ [by Theorems 2.3.3 and 2.3.9 in 10], the larger $|\rho_j|$ is, the faster $|A_j|$ converges to 0 and the smaller the error in the approximation of the right tail of $U_j$ by a purely Pareto tail will be. Any distribution part of the Hall-Welsh class (5) satisfies this condition [by Theorem 2.3.9 in 10].

The next theoretical contribution of this paper, of interest in its own right, examines the joint convergence between Hill estimators and intermediate order statistics across marginals. The proof essentially relies on asymptotic representations of the $\widehat{\gamma}_{\tau_n,j}$ and $\widehat{q}_{\tau_n,j}$ as sums of independent random variables constructed on the $X_{i,j}$ $(1 \le i \le n)$ in a framework of average excesses as defined in [33], with the desired joint asymptotic normality then following from the Cramér-Wold device.

**Theorem 2.2.** *Assume that Conditions A and B hold. Let $\tau_n \uparrow 1$ be such that $n(1-\tau_n) \to \infty$ and, for any $1 \le j \le d$, $\sqrt{n(1-\tau_n)}A_j((1-\tau_n)^{-1}) \to \lambda_j \in \mathbb{R}$ as $n \to \infty$. Then we have*

$$\sqrt{n(1-\tau_n)}\left(\widehat{\gamma}_{\tau_n,j} - \gamma_j, \frac{\widehat{q}_{\tau_n,j}}{q_{\tau_n,j}} - 1\right)_{1\le j\le d} \xrightarrow{\mathrm{d}} \mathcal{N}_{2d}\left((\lambda_j/(1-\rho_j), 0)_{1\le j\le d}, \mathbf{\Sigma}^{\mathrm{Q}}(\boldsymbol{\gamma}, \boldsymbol{R})\right).$$

*The covariance matrix $\mathbf{\Sigma}^{\mathrm{Q}}(\boldsymbol{\gamma}, \boldsymbol{R})$ can be partitioned into $d^2$ square blocks $\mathbf{\Sigma}^{\mathrm{Q}}_{j,\ell}(\boldsymbol{\gamma}, \boldsymbol{R})$, with the diagonal blocks given by $\mathbf{\Sigma}^{\mathrm{Q}}_{j,j}(\boldsymbol{\gamma}, \boldsymbol{R}) = \gamma_j^2 \mathrm{I}_2$ (where $\mathrm{I}_2$ denotes the $2\times 2$ identity matrix) for any*

$j \in \{1, \ldots, d\}$, *and the off-diagonal blocks given by*

$$\boldsymbol{\Sigma}^{\mathrm{Q}}_{j,\ell}(\boldsymbol{\gamma}, \boldsymbol{R}) = \gamma_j \gamma_\ell \begin{pmatrix} R_{j,\ell}(1,1) & \displaystyle\int_0^1 R_{j,\ell}(u,1) \frac{\mathrm{d}u}{u} - R_{j,\ell}(1,1) \\ \displaystyle\int_0^1 R_{j,\ell}(1,u) \frac{\mathrm{d}u}{u} - R_{j,\ell}(1,1) & R_{j,\ell}(1,1) \end{pmatrix}$$

*for any* $j, \ell \in \{1, \ldots, d\}$ *with* $j < \ell$.

Theorem 2.2 is written in our tail dependence framework (through Condition A) using the standard conditions for the analysis of the Hill estimators $\widehat{\gamma}_{\tau_n,j}$ (these are: Condition B and the related bias conditions $\sqrt{n(1-\tau_n)} A_j((1-\tau_n)^{-1}) \to \lambda_j \in \mathbb{R}$ as $n \to \infty$, for $1 \le j \le d$). Note that there are generally nonzero correlations (controlled by the tail copulae) between pairs of Hill estimators, pairs of intermediate order statistics, as well as between the Hill estimator of a given marginal and an intermediate order statistic pertaining to another marginal.

The desired result on the joint convergence of the $\widehat{\xi}_{\tau_n,j}$ is now a corollary of Theorem 2.2 and the delta-method. Set $m(x) = (1-x)^{-1} - \log(x^{-1} - 1)$, for $x \in (0,1)$.

**Corollary 2.3.** *Work under the conditions of Theorem 2.2. Assume in addition that* $\mathbb{E}|\min(X_j, 0)| < \infty$, *that* $0 < \gamma_j < 1$ *and that* $\sqrt{n(1-\tau_n)} q_{\tau_n,j}^{-1} \to \mu_j \in \mathbb{R}$ *as* $n \to \infty$ *for any* $1 \le j \le d$. *Then*

$$\sqrt{n(1-\tau_n)} \left( \frac{\widehat{\xi}_{\tau_n,j}}{\xi_{\tau_n,j}} - 1 \right)_{1 \le j \le d} \xrightarrow{\mathrm{d}} \mathcal{N}_d \left( \boldsymbol{b}, \boldsymbol{V}^{\mathrm{QB}}(\boldsymbol{\gamma}, \boldsymbol{R}) \right),$$

*where the asymptotic bias* $\boldsymbol{b}$ *and the covariance matrix* $\boldsymbol{V}^{\mathrm{QB}}(\boldsymbol{\gamma}, \boldsymbol{R})$ *have entries*

$$\boldsymbol{b}_j = -\gamma_j (\gamma_j^{-1} - 1)^{\gamma_j} \mathbb{E}(X_j) \mu_j + \left( \frac{m(\gamma_j)}{1-\rho_j} - \frac{(\gamma_j^{-1}-1)^{-\rho_j}}{1-\gamma_j-\rho_j} - \frac{(\gamma_j^{-1}-1)^{-\rho_j} - 1}{\rho_j} \right) \lambda_j$$

$$\textit{and } \boldsymbol{V}^{\mathrm{QB}}_{j,\ell}(\boldsymbol{\gamma}, \boldsymbol{R}) = \begin{cases} \gamma_j^2 (1 + (m(\gamma_j))^2) & \text{if } j = \ell, \\ \gamma_j \gamma_\ell \Bigg( R_{j,\ell}(1,1)(m(\gamma_j) - 1)(m(\gamma_\ell) - 1) & \\ \quad + m(\gamma_j) \displaystyle\int_0^1 R_{j,\ell}(u,1) \frac{\mathrm{d}u}{u} + m(\gamma_\ell) \int_0^1 R_{j,\ell}(1,u) \frac{\mathrm{d}u}{u} \Bigg) & \text{if } j < \ell. \end{cases}$$

This result is the multivariate extension of Corollary 2 in [7] that is required for our purposes. Unlike the latter, Corollary 2.3 is formulated without the unnecessary assumption of an increasing (marginal) distribution function. The integrability conditions $\mathbb{E}|\min(X_j, 0)| < \infty$ and $0 < \gamma_j < 1$ for any $1 \le j \le d$ are weaker than the corresponding assumptions in Theorem 2.1: this is because the validity of approximation (4) only requires that expectiles exist, which is equivalent to assuming a finite first moment of the $X_j$. Note also that compared to Theorem 2.1, Corollary 2.3 features the additional assumption that $\sqrt{n(1-\tau_n)} q_{\tau_n,j}^{-1} \to \mu_j \in \mathbb{R}$ as $n \to \infty$. This is due to the presence of a bias term proportional to the reciprocal of $q_{\tau_n,j}$ in the remainder term of approximation (4) (see Proposition 1(i) in [9]) whose use is instrumental in the construction of $\widehat{\xi}_{\tau_n,j}$. Since the functions $t \mapsto 1/q_{1-t^{-1},j}$ and $t \mapsto |A_j(t)|$ are respectively regularly varying with indices $-\gamma_j$ and $\rho_j$, one will have $\mu_j = 0$ for all $j$ if $\max_j \rho_j > \max_j(-\gamma_j)$. On the contrary, if $\max_j \rho_j < \max_j(-\gamma_j)$, then $\lambda_j = 0$ for all $j$.

## 2.2. At the extreme level

We consider now the problem of most relevance to risk management in practice, which is to estimate extreme expectiles $\xi_{\tau_n',j}$, where $\tau_n' \to 1$ is such that $n(1-\tau_n') \to c \in [0,\infty)$. In risk management, one would typically consider $\tau_n' \geq 1 - 1/n$, see for example Chapter 4 of [10] and [6] in the context of quantile-based extreme value estimation. The basic idea is to extrapolate intermediate expectile estimators at level $\tau_n$ to the extreme level $\tau_n'$, beyond the observed data, using the marginal heavy tails assumption. This is warranted by convergence (4), which entails

$$\frac{\xi_{\tau_n',j}}{\xi_{\tau_n,j}} \approx \frac{q_{\tau_n',j}}{q_{\tau_n,j}} = \frac{U_j((1-\tau_n')^{-1})}{U_j((1-\tau_n)^{-1})} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma_j} \quad \text{as } n \to \infty.$$

This suggests the following two extrapolating estimators:

$$\widetilde{\xi}^{\star}_{\tau_n',j} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}_{\tau_n,j}} \widetilde{\xi}_{\tau_n,j} \quad \text{(LAWS-based extrapolating estimator)}$$

$$\widehat{\xi}^{\star}_{\tau_n',j} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}_{\tau_n,j}} \widehat{\xi}_{\tau_n,j} = (\widehat{\gamma}^{-1}_{\tau_n,j} - 1)^{-\widehat{\gamma}_{\tau_n,j}} \widehat{q}^{\star}_{\tau_n',j} \quad \text{(QB extrapolating estimator)}$$

where $\widehat{q}^{\star}_{\tau_n',j}$ is the Weissman estimator of the extreme quantile $q_{\tau_n',j}$ (see [35]) and $\tau_n$ is an anchor intermediate level specified by the user. The next main result towards our goal of carrying out joint inference about multiple extreme expectiles is a statement of the joint convergence of these estimators across marginals. The key to the proof of this result is to note that any estimator of the form

$$\overline{\xi}^{\star}_{\tau_n',j} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}_{\tau_n,j}} \overline{\xi}_{\tau_n,j}, \tag{6}$$

where $\overline{\xi}_{\tau_n,j}$ is a $\sqrt{n(1-\tau_n)}$−consistent estimator of $\xi_{\tau_n,j}$, satisfies

$$\log\left(\frac{\overline{\xi}^{\star}_{\tau_n',j}}{\xi_{\tau_n',j}}\right) = (\widehat{\gamma}_{\tau_n,j} - \gamma_j)\log\left(\frac{1-\tau_n}{1-\tau_n'}\right) + \log\left(\frac{\overline{\xi}_{\tau_n,j}}{\xi_{\tau_n,j}}\right) - \log\left(\left[\frac{1-\tau_n'}{1-\tau_n}\right]^{\gamma_j} \frac{\xi_{\tau_n',j}}{\xi_{\tau_n,j}}\right). \tag{7}$$

Under suitable conditions, the second (random) term and the third (bias) term are dominated by the first term linked to the Hill estimator, leading to the common asymptotic distribution obtained below.

**Theorem 2.4.** *Assume that Conditions A and B hold, with $\rho_j < 0$ for any $1 \leq j \leq d$. Let $\tau_n, \tau_n' \uparrow 1$ with $n(1-\tau_n) \to \infty$, $n(1-\tau_n') \to c \in [0,\infty)$ and $\sqrt{n(1-\tau_n)}/\log[(1-\tau_n)/(1-\tau_n')] \to \infty$ as $n \to \infty$. Assume also that for any $1 \leq j \leq d$, $\sqrt{n(1-\tau_n)}q^{-1}_{\tau_n,j} \to \mu_j \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}A_j((1-\tau_n)^{-1}) \to \lambda_j \in \mathbb{R}$ as $n \to \infty$. Let $\boldsymbol{b}^\star = (\lambda_j/(1-\rho_j))_{1\leq j\leq d}$ and define a covariance matrix $\boldsymbol{V}^\star(\boldsymbol{\gamma},\boldsymbol{R})$ by*

$$\boldsymbol{V}^\star_{j,\ell}(\boldsymbol{\gamma},\boldsymbol{R}) = \begin{cases} \gamma_j^2 & \text{if } j = \ell, \\ \gamma_j\gamma_\ell R_{j,\ell}(1,1) & \text{if } j < \ell. \end{cases}$$

*(i) Assume that there is $\delta > 0$ such that $\mathbb{E}|\min(X_j,0)|^{2+\delta} < \infty$ and that $0 < \gamma_j < 1/2$ for any $1 \le j \le d$. Then*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widetilde{\xi}^{\star}_{\tau_n',j}}{\xi_{\tau_n',j}} - 1\right)_{1\le j\le d} \xrightarrow{\mathrm{d}} \mathcal{N}_d\left(\boldsymbol{b}^{\star}, \boldsymbol{V}^{\star}(\boldsymbol{\gamma},\boldsymbol{R})\right).$$

*(ii) Assume that $\mathbb{E}|\min(X_j,0)| < \infty$ and that $0 < \gamma_j < 1$ for any $1 \le j \le d$. Then*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widehat{\xi}^{\star}_{\tau_n',j}}{\xi_{\tau_n',j}} - 1\right)_{1\le j\le d} \xrightarrow{\mathrm{d}} \mathcal{N}_d\left(\boldsymbol{b}^{\star}, \boldsymbol{V}^{\star}(\boldsymbol{\gamma},\boldsymbol{R})\right).$$

This result generalises Corollaries 3 and 4 of [7] to the joint inference setting. However, even though the asymptotics in Theorem 2.4 are dominated by those of the Hill estimators, correctly inferring the anchor intermediate expectile will also be important in finite-sample situations, as we shall show in our construction of confidence regions and in our simulation study.

# 3. Joint inference on extreme expectiles

Equipped with the theory developed in Section 2, we derive asymptotic confidence regions for inference about extreme expectiles and provide a testing procedure for their equality. We start by the construction of confidence regions at intermediate and extreme levels. Treating the intermediate case first will be a key step for the definition of accurate Gaussian confidence regions for multiple extreme expectiles. Throughout this section, we let $\boldsymbol{\xi}_{\tau_n} = (\xi_{\tau_n,1},\ldots,\xi_{\tau_n,d})^{\top}$ and define similarly $\boldsymbol{\xi}_{\tau_n'}$, $\widetilde{\boldsymbol{\xi}}_{\tau_n}$, $\widehat{\boldsymbol{\xi}}_{\tau_n}$, $\widetilde{\boldsymbol{\xi}}^{\star}_{\tau_n'}$ and $\widehat{\boldsymbol{\xi}}^{\star}_{\tau_n'}$. The symbol $\mathbf{1}_d$ denotes the $d-$dimensional vector with all entries equal to 1. All operations on vectors, apart from matrix operations, are meant componentwise.

## 3.1. Asymptotic confidence region construction: intermediate case

To save space, we only highlight the key steps of our construction of intermediate expectile confidence regions; an expanded discussion can be found in Appendix B of the supplementary file.

**Using LAWS estimation** The main instrument is Theorem 2.1, namely

$$\sqrt{n(1-\tau_n)}\left(\frac{\widetilde{\boldsymbol{\xi}}_{\tau_n}}{\boldsymbol{\xi}_{\tau_n}} - \mathbf{1}_d\right) \xrightarrow{\mathrm{d}} \mathcal{N}_d(\mathbf{0}_d, \boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})). \tag{8}$$

Using this Gaussian asymptotic approximation to build a confidence region for $\boldsymbol{\xi}_{\tau_n}$ is a delicate task; we show in Table I of Appendix B how a standard Gaussian confidence interval for a marginal intermediate expectile has an effective coverage probability generally close to twice the nominal level.

We investigate a solution based on the proof of Theorem 2.1. Let $\varphi_\tau(y) = |\tau - \mathbb{1}\{y \le 0\}|\, y$ be the derivative of $\eta_\tau/2$. A careful analysis of this proof and extensive Monte-Carlo simulations suggest that $\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ can be well approximated with

$$\boldsymbol{V}^{\mathrm{LAWS}}_{j,j}(\boldsymbol{\gamma},\boldsymbol{R}) \approx \frac{2\gamma_j^2}{1-2\gamma_j} \times \frac{1+\overline{F}_j(\xi_{\tau_n,j})/(1-\tau_n)}{[1+(2\tau_n-1)\overline{F}_j(\xi_{\tau_n,j})/(1-\tau_n)]^2} \quad \text{for } j=\ell$$

$$\text{and } \boldsymbol{V}^{\mathrm{LAWS}}_{j,\ell}(\boldsymbol{\gamma},\boldsymbol{R}) \approx \gamma_j\gamma_\ell \frac{\mathbb{E}(\varphi_{\tau_n}(X_j-\xi_{\tau_n,j})\varphi_{\tau_n}(X_\ell-\xi_{\tau_n,\ell}))}{(1-\tau_n)\xi_{\tau_n,j}\xi_{\tau_n,\ell}} \text{ for } j\neq\ell$$

as $n\to\infty$. Our estimator of $\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ is constructed by plugging in the LAWS and Hill estimators, the empirical survival functions $\widehat{\overline{F}}_{n,j}$ based on the $X_{i,j}$ $(1\le i\le n)$ and the empirical covariances

$$\overline{m}_{n,j,\ell} = \frac{1}{n}\sum_{i=1}^{n}\varphi_{\tau_n}(X_{i,j}-\widetilde{\xi}_{\tau_n,j})\varphi_{\tau_n}(X_{i,\ell}-\widetilde{\xi}_{\tau_n,\ell}).$$

This results in the estimator $\widehat{\boldsymbol{V}}^{\mathrm{LAWS}}_n(\boldsymbol{\gamma},\boldsymbol{R})$ of $\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ given elementwise by

$$\widehat{\boldsymbol{V}}^{\mathrm{LAWS}}_{n,j,j}(\boldsymbol{\gamma},\boldsymbol{R}) = \frac{2\widehat{\gamma}^2_{\tau_n,j}}{1-2\widehat{\gamma}_{\tau_n,j}} \times \frac{1+\widehat{\overline{F}}_{n,j}(\widetilde{\xi}_{\tau_n,j})/(1-\tau_n)}{\left[1+(2\tau_n-1)\widehat{\overline{F}}_{n,j}(\widetilde{\xi}_{\tau_n,j})/(1-\tau_n)\right]^2} \text{ for } j=\ell$$

$$\text{and } \widehat{\boldsymbol{V}}^{\mathrm{LAWS}}_{n,j,\ell}(\boldsymbol{\gamma},\boldsymbol{R}) = \widehat{\gamma}_{\tau_n,j}\widehat{\gamma}_{\tau_n,\ell}\frac{\overline{m}_{n,j,\ell}}{(1-\tau_n)\widetilde{\xi}_{\tau_n,j}\widetilde{\xi}_{\tau_n,\ell}} \text{ for } j\neq\ell.$$

Under the assumptions of Theorem 2.1, Proposition A.4 in the Appendix shows that this is indeed a consistent estimator of $\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$. When $\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ is symmetric positive definite, multiplying the left-hand side in (8) by the positive definite inverse square root $[\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})]^{-1/2}$ of $\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ and then plugging in the estimator $\widehat{\boldsymbol{V}}^{\mathrm{LAWS}}_n(\boldsymbol{\gamma},\boldsymbol{R})$ produces an asymptotically Gaussian random vector with independent standard Gaussian components. Therefore, if $\chi^2_{d,1-\alpha}$ denotes the $(1-\alpha)-$quantile of the chi-square distribution with $d$ degrees of freedom, and $B_d(\mathbf{0}_d,r)$ denotes the closed Euclidean ball in $\mathbb{R}^d$ whose centre is the origin $\mathbf{0}_d$ and radius is $r$, we find an $(1-\alpha)-$asymptotic LAWS-based confidence region for $\boldsymbol{\xi}_{\tau_n}$ as the random ellipsoid

$$\begin{aligned}\widetilde{\mathcal{E}}_{\tau_n,\alpha} &= \widetilde{\boldsymbol{\xi}}_{\tau_n}\left[\mathbf{1}_d + \left[\widehat{\boldsymbol{V}}^{\mathrm{LAWS}}_n(\boldsymbol{\gamma},\boldsymbol{R})\right]^{1/2} B_d\left(\mathbf{0}_d, \sqrt{\chi^2_{d,1-\alpha}/n(1-\tau_n)}\right)\right] \\ &= \left\{\mathbf{z}\in\mathbb{R}^d \,|\, \exists \mathbf{u}\in B_d\left(\mathbf{0}_d, \sqrt{\chi^2_{d,1-\alpha}/n(1-\tau_n)}\right),\ \mathbf{z} = \widetilde{\boldsymbol{\xi}}_{\tau_n}\left[\mathbf{1}_d + \left[\widehat{\boldsymbol{V}}^{\mathrm{LAWS}}_n(\boldsymbol{\gamma},\boldsymbol{R})\right]^{1/2}\mathbf{u}\right]\right\}.\end{aligned}$$

[Recall that all operations except the matrix product $\left[\widehat{\boldsymbol{V}}^{\mathrm{LAWS}}_n(\boldsymbol{\gamma},\boldsymbol{R})\right]^{1/2}\mathbf{u}$ are meant componentwise.] This construction is illustrated in the top panel of Figure I (see Appendix B) and formalised in the following result.

**Proposition 3.1.** *Work under the conditions of Theorem 2.1. If $\boldsymbol{V}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ is positive definite, then $\mathbb{P}(\boldsymbol{\xi}_{\tau_n}\in\widetilde{\mathcal{E}}_{\tau_n,\alpha})\to 1-\alpha$ as $n\to\infty$.*

**Using QB estimation** With the QB estimator, the main tool is Corollary 2.3. Similarly to what is observed when using LAWS estimators, great care has to be taken in constructing asymptotic confidence regions. Table I in Appendix B reports the coverage probability of a straightforward Gaussian confidence interval for a marginal expectile based on Corollary 2.3; this is often a very poor confidence interval whose coverage probability is much lower than the nominal level.

Contrary to the LAWS estimator, the QB estimator is asymptotically biased due to its reliance on the relationship (4). The $j$th component of this bias is essentially

$$\boldsymbol{b}_j \approx -\gamma_j(\gamma_j^{-1}-1)^{\gamma_j}\mathbb{E}(X_j)\frac{\sqrt{n(1-\tau_n)}}{q_{\tau_n,j}} + \left(\frac{m(\gamma_j)}{1-\rho_j} - \frac{(\gamma_j^{-1}-1)^{-\rho_j}}{1-\gamma_j-\rho_j} - \frac{(\gamma_j^{-1}-1)^{-\rho_j}-1}{\rho_j}\right)\sqrt{n(1-\tau_n)}A_j((1-\tau_n)^{-1}) \text{ as } n\to\infty.$$

Two sources of bias therefore arise when using the QB estimator: one proportional to $1/q_{\tau_n,j}$ (representing marginal tail heaviness), and the other to $A_j((1-\tau_n)^{-1})$ (representing second-order behaviour). The correction of the latter source of bias involves estimating accurately the second-order parameter $\rho_j$, which is a notoriously difficult problem [see *e.g.* the Introduction of 5], especially from the practical point of view since consistent estimators of $\rho_j$ typically suffer from low rates of convergence, see *e.g.* [18, p.2638] and [19, p.298]. As such, correcting second-order bias tends to increase finite-sample variability substantially, resulting in confidence regions that may be too conservative. By contrast, the simple expression of the bias component proportional to $1/q_{\tau_n,j}$ makes its correction a straightforward task, with all estimators involved converging at the rate $\sqrt{n(1-\tau_n)}$ or more. This constitutes our rationale for concentrating specifically on the first source of bias with the estimator

$$\widehat{\boldsymbol{b}}_j = -\widehat{\gamma}_{\tau_n,j}(\widehat{\gamma}_{\tau_n,j}^{-1}-1)^{\widehat{\gamma}_{\tau_n,j}}\overline{X}_{n,j}\frac{\sqrt{n(1-\tau_n)}}{\widehat{q}_{\tau_n,j}}, \text{ where } \overline{X}_{n,j} = \frac{1}{n}\sum_{i=1}^n X_{i,j}.$$

This is justified from the theoretical point of view as soon as $|\rho_j| > \gamma_j$ because then $\lambda_j = 0$, see the discussion below Corollary 2.3; our experience in finite samples is that correcting this source of bias alone often brings a very substantial improvement, regardless of the respective positions of $|\rho_j|$ and $\gamma_j$ (see Appendix B in the Supplementary Material document for simple examples where this is numerically justified). The covariance matrix $\boldsymbol{V}^{\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R})$, meanwhile, may be estimated as follows:

$$\widehat{\boldsymbol{V}}^{\mathrm{QB}}_{n,j,j}(\boldsymbol{\gamma},\boldsymbol{R}) = \widehat{\gamma}^2_{\tau_n,j}(1+(m(\widehat{\gamma}_{\tau_n,j}))^2) \text{ (with } m(x)=(1-x)^{-1}-\log(x^{-1}-1))$$

$$\text{and } \widehat{\boldsymbol{V}}^{\mathrm{QB}}_{n,j,\ell}(\boldsymbol{\gamma},\boldsymbol{R}) = \widehat{\gamma}_{\tau_n,j}\widehat{\gamma}_{\tau_n,\ell}\Bigg(\widehat{R}_{\tau_n,j,\ell}(1,1)(m(\widehat{\gamma}_{\tau_n,j})-1)(m(\widehat{\gamma}_{\tau_n,\ell})-1) + m(\widehat{\gamma}_{\tau_n,j})\int_0^1 \widehat{R}_{\tau_n,j,\ell}(u,1)\frac{\mathrm{d}u}{u} + m(\widehat{\gamma}_{\tau_n,\ell})\int_0^1 \widehat{R}_{\tau_n,j,\ell}(1,u)\frac{\mathrm{d}u}{u}\Bigg)$$

where the estimator of the tail copula function $R_{j,\ell}$ is defined as

$$\widehat{R}_{\tau_n,j,\ell}(u,v) = \frac{1}{n(1-\tau_n)}\sum_{i=1}^n \mathbb{1}\left\{\frac{n+1-r_{n,i,j}}{(n+1)(1-\tau_n)} \le u, \frac{n+1-r_{n,i,\ell}}{(n+1)(1-\tau_n)} \le v\right\}. \tag{9}$$

[Here $r_{n,i,j}$ denotes the marginal rank of observation $X_{i,j}$ among $X_{1,j}, X_{2,j}, \ldots, X_{n,j}$.] This estimator is a slightly modified version of the estimator of the empirical upper tail copula estimator given in Equation (13) in [30]. The estimator $\widehat{\boldsymbol{V}}^{\mathrm{QB}}_n(\boldsymbol{\gamma},\boldsymbol{R})$ is a consistent estimator of $\boldsymbol{V}^{\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R})$ and therefore an $(1-\alpha)-$asymptotic QB confidence region for $\boldsymbol{\xi}_{\tau_n}$ is the random ellipsoid

$$\widehat{\mathcal{E}}_{\tau_n,\alpha} = \widehat{\boldsymbol{\xi}}_{\tau_n}\left[\mathbf{1}_d - \frac{\widehat{\boldsymbol{b}}}{\sqrt{n(1-\tau_n)}} + \left[\widehat{\boldsymbol{V}}^{\mathrm{QB}}_n(\boldsymbol{\gamma},\boldsymbol{R})\right]^{1/2} B_d\left(\mathbf{0}_d, \sqrt{\chi^2_{d,1-\alpha}/n(1-\tau_n)}\right)\right].$$

We formalise this result as follows.

**Proposition 3.2.** *Work under the conditions of Corollary 2.3, with additionally $\lambda_j = 0$ for all $j$. If $\boldsymbol{V}^{\mathrm{QB}}(\boldsymbol{\gamma}, \boldsymbol{R})$ is positive definite, then $\mathbb{P}(\boldsymbol{\xi}_{\tau_n} \in \widehat{\mathcal{E}}_{\tau_n,\alpha}) \to 1-\alpha$ as $n \to \infty$.*

The QB confidence region construction is illustrated in the top panel of Figure I in Appendix B, where it is seen that the QB region tends to have a lower volume than the LAWS region.

## 3.2. Asymptotic confidence region construction: extreme case

At the extreme level, the key result for our purposes is Theorem 2.4. Nevertheless, if one constructs an asymptotic confidence region directly from this result, the actual finite-sample coverage probability can be quite poor, even in the estimation of a single extreme expectile: see Appendix A.2 in [7] where Gaussian QQ-plots show that the observed variance of extreme expectile estimators can be fairly different from the asymptotic variance in the Gaussian approximation. We illustrate this in more detail in Appendix C.

Our idea is to, first, get a finer understanding of the uncertainty in the estimation of extreme expectiles. The gist of our method is that, even though a common asymptotic distribution is found for extreme expectile estimators of the form (6) regardless of the actual intermediate expectile estimator chosen, in practice, the behaviour of $\overline{\xi}_{\tau_n,j}$ matters, and so does the correlation between $\overline{\xi}_{\tau_n,j}$ and $\widehat{\gamma}_{\tau_n,j}$. This is especially true when $\log d_n = \log[(1-\tau_n)/(1-\tau_n')]$, representing the cost of the extrapolation methodology, is only moderately large. Investigating this uncertainty and correlation will lead us to define corrected Gaussian asymptotic confidence regions. Each confidence region will be constructed on the log-scale; using this scale has been shown to improve finite-sample coverage of confidence regions for extreme risk measures [see *e.g.* p.628 in 13, in the context of extreme quantile estimation]. We found from Monte-Carlo simulations that this is also the case for expectiles.

**Using the LAWS-based extrapolating estimator** The crucial result is an extension of Theorem 2.1 giving the joint convergence of the Hill estimators $\widehat{\gamma}_{\tau_n,j}$ and intermediate LAWS expectile estimators $\widetilde{\xi}_{\tau_n,j}$ across marginals. This joint convergence is proven by putting them in a common (artificial) minimisation framework: namely, we note that

$$\sqrt{n(1-\tau_n)}\left(\widehat{\gamma}_{\tau_n,j}-\gamma_j\right)_{1\le j\le d} = \underset{\mathbf{u}\in\mathbb{R}^d}{\arg\min} \sum_{j=1}^{d} \frac{1}{2}\left(u_j - \sqrt{n(1-\tau_n)}(\widehat{\gamma}_{\tau_n,j}-\gamma_j)\right)^2.$$

Combined with (3) and the asymptotic representation of the $\widehat{\gamma}_{\tau_n,j}$ as sums of independent random variables constructed on the $X_{i,j}$ that is already used in the proof of Theorem 2.2, this makes it possible to view Hill estimators and intermediate LAWS expectile estimators jointly as minimisers of a convex random function, whose convergence entails the required joint convergence by Theorem 5 in [25].

**Theorem 3.3.** *Assume that Conditions A and B hold. Assume further that there is $\delta > 0$ such that $\mathbb{E}|\min(X_j,0)|^{2+\delta} < \infty$ and that $0 < \gamma_j < 1/2$ for any $1 \le j \le d$. Let $\tau_n \uparrow 1$ be such that $n(1-\tau_n) \to \infty$ and, for any $1 \le j \le d$, $\sqrt{n(1-\tau_n)}A_j((1-\tau_n)^{-1}) \to \lambda_j \in \mathbb{R}$ as $n \to \infty$. Then we have*

$$\sqrt{n(1-\tau_n)}\left(\widehat{\gamma}_{\tau_n,j}-\gamma_j, \frac{\widetilde{\xi}_{\tau_n,j}}{\xi_{\tau_n,j}}-1\right)_{1\le j\le d} \xrightarrow{\mathrm{d}} \mathcal{N}_{2d}\left((\lambda_j/(1-\rho_j),0)_{1\le j\le d}, \boldsymbol{\Sigma}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})\right).$$

*The covariance matrix* $\boldsymbol{\Sigma}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ *is partitioned into* $d^2$ *square blocks* $\boldsymbol{\Sigma}^{\mathrm{LAWS}}_{j,\ell}(\boldsymbol{\gamma},\boldsymbol{R})$ *given by*

$$\boldsymbol{\Sigma}^{\mathrm{LAWS}}_{j,j}(\boldsymbol{\gamma},\boldsymbol{R}) = \gamma_j^2 \begin{pmatrix} 1 & \dfrac{\gamma_j(\gamma_j^{-1}-1)^{\gamma_j}}{(1-\gamma_j)^2} \\ \dfrac{\gamma_j(\gamma_j^{-1}-1)^{\gamma_j}}{(1-\gamma_j)^2} & \dfrac{2\gamma_j}{1-2\gamma_j} \end{pmatrix}$$

*when* $j=\ell\in\{1,\ldots,d\}$ *and, elementwise,*

$$\begin{aligned}
\boldsymbol{\Sigma}^{\mathrm{LAWS}}_{j,\ell}(\boldsymbol{\gamma},\boldsymbol{R})(1,1) &= \gamma_j\gamma_\ell R_{j,\ell}(1,1),\\
\boldsymbol{\Sigma}^{\mathrm{LAWS}}_{j,\ell}(\boldsymbol{\gamma},\boldsymbol{R})(1,2) &= \gamma_\ell \iint_{[1,\infty)^2} R_{j,\ell}\left(x_j^{-1/\gamma_j},(\gamma_\ell^{-1}-1)x_\ell^{-1/\gamma_\ell}\right)\frac{\mathrm{d}x_j}{x_j}\,\mathrm{d}x_\ell\\
&\quad -\gamma_j\gamma_\ell\int_1^\infty R_{j,\ell}\left(1,(\gamma_\ell^{-1}-1)x_\ell^{-1/\gamma_\ell}\right)\mathrm{d}x_\ell,\\
\boldsymbol{\Sigma}^{\mathrm{LAWS}}_{j,\ell}(\boldsymbol{\gamma},\boldsymbol{R})(2,1) &= \gamma_j \iint_{[1,\infty)^2} R_{j,\ell}\left((\gamma_j^{-1}-1)x_j^{-1/\gamma_j},x_\ell^{-1/\gamma_\ell}\right)\mathrm{d}x_j\,\frac{\mathrm{d}x_\ell}{x_\ell}\\
&\quad -\gamma_j\gamma_\ell\int_1^\infty R_{j,\ell}\left((\gamma_j^{-1}-1)x_j^{-1/\gamma_j},1\right)\mathrm{d}x_j,\\
\boldsymbol{\Sigma}^{\mathrm{LAWS}}_{j,\ell}(\boldsymbol{\gamma},\boldsymbol{R})(2,2) &= \gamma_j\gamma_\ell \iint_{[1,\infty)^2} R_{j,\ell}\left((\gamma_j^{-1}-1)x_j^{-1/\gamma_j},(\gamma_\ell^{-1}-1)x_\ell^{-1/\gamma_\ell}\right)\mathrm{d}x_j\,\mathrm{d}x_\ell
\end{aligned}$$

*for any* $j,\ell\in\{1,\ldots,d\}$ *with* $j<\ell$.

Theorem 3.3 and Equation (7) suggest the following approximation for the LAWS-based extrapolating estimator, as $n\to\infty$, on the log-scale:

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left[\log\left(\frac{\widetilde{\boldsymbol{\xi}}^{\star}_{\tau_n'}}{\boldsymbol{\xi}_{\tau_n'}}\right)+\log\left(\left[\frac{1-\tau_n'}{1-\tau_n}\right]^{\boldsymbol{\gamma}}\frac{\boldsymbol{\xi}_{\tau_n'}}{\boldsymbol{\xi}_{\tau_n}}\right)\right]\approx\mathcal{N}_d\left(\mathbf{0}_d,\boldsymbol{V}_n^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})\right), \tag{10}$$

where $\boldsymbol{V}_n^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ is defined elementwise as (recall that $\log d_n=\log[(1-\tau_n)/(1-\tau_n')]$)

$$\boldsymbol{V}^{\star,\mathrm{LAWS}}_{n,j,\ell}(\boldsymbol{\gamma},\boldsymbol{R}) = \begin{pmatrix}1\\ 1/\log d_n\end{pmatrix}^{\top}\boldsymbol{\Sigma}^{\mathrm{LAWS}}_{j,\ell}(\boldsymbol{\gamma},\boldsymbol{R})\begin{pmatrix}1\\ 1/\log d_n\end{pmatrix}.$$

We now focus on the estimation of the bias term appearing in approximation (10), and of the matrix $\boldsymbol{V}_n^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$. Use Proposition 1(i) in [9] and the proof of Theorem 4.3.8 in [10] to find, as $n\to\infty$,

$$-\log\left(\left[\frac{1-\tau_n'}{1-\tau_n}\right]^{\gamma_j}\frac{\xi_{\tau_n',j}}{\xi_{\tau_n,j}}\right)\approx\frac{\gamma_j(\gamma_j^{-1}-1)^{\gamma_j}\mathbb{E}(X_j)}{q_{\tau_n,j}}+\mathrm{O}(A_j((1-\tau_n)^{-1})). \tag{11}$$

Here and as above we neither emphasise nor estimate the bias term proportional to $A_j((1-\tau_n)^{-1})$. We therefore estimate the left-hand side above by $-\widehat{\boldsymbol{b}}_j/\sqrt{n(1-\tau_n)}$ (see Section 3.1). To find an estimator of the covariance matrix $\boldsymbol{V}_n^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$, we note that $\boldsymbol{\Sigma}_{j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(1,1)$ can be estimated by

$$\widehat{\boldsymbol{\Sigma}}_{n,j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(1,1) = \begin{cases} \widehat{\gamma}_{\tau_n,j}^2 & \text{if } j=\ell \\ \widehat{\gamma}_{\tau_n,j}\widehat{\gamma}_{\tau_n,\ell}\widehat{R}_{\tau_n,j,\ell}(1,1) & \text{if } j<\ell \end{cases}$$

with the notation of Section 3.1. Similarly $\boldsymbol{\Sigma}_{j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(2,2) = \boldsymbol{V}_{j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ is estimated with $\widehat{\boldsymbol{V}}_{n,j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$. An estimator for the off-diagonal entry $\boldsymbol{\Sigma}_{j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(1,2)$ is obtained by recalling that (see the proof of Theorem 3.3 in Appendix A):

$$\begin{aligned}\boldsymbol{\Sigma}_{j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(1,2) &\approx \gamma_j\gamma_\ell \frac{\operatorname{Cov}\left([\log X_j - \log q_{\tau_n,j}]\mathbb{1}\{X_j > q_{\tau_n,j}\}, \varphi_{\tau_n}(X_\ell - \xi_{\tau_n,\ell})\right)}{\mathbb{E}([\log X_j - \log q_{\tau_n,j}]\mathbb{1}\{X_j > q_{\tau_n,j}\})\xi_{\tau_n,\ell}} \\ &\quad - \gamma_j\gamma_\ell \frac{\operatorname{Cov}\left(\mathbb{1}\{X_j > q_{\tau_n,j}\}, \varphi_{\tau_n}(X_\ell - \xi_{\tau_n,\ell})\right)}{\mathbb{P}(X_j > q_{\tau_n,j})\xi_{\tau_n,\ell}}, \quad \text{as } n\to\infty.\end{aligned}$$

We thus estimate $\boldsymbol{\Sigma}_{j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(1,2)$ with $\widehat{\boldsymbol{\Sigma}}_{n,j,j}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(1,2) = \widehat{\gamma}_{\tau_n,j}^3(\widehat{\gamma}_{\tau_n,j}^{-1}-1)^{\widehat{\gamma}_{\tau_n,j}}/(1-\widehat{\gamma}_{\tau_n,j})^2$ when $j=\ell$, and otherwise by (recall that $\varphi_{\tau_n}(X_\ell - \xi_{\tau_n,\ell})$ has expectation 0):

$$\begin{aligned}&\widehat{\boldsymbol{\Sigma}}_{n,j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(1,2) \\ &= \widehat{\gamma}_{\tau_n,\ell}\cdot\frac{1}{n}\sum_{i=1}^n [\log X_{i,j} - \log X_{n-\lfloor n(1-\tau_n)\rfloor,n,j}]\mathbb{1}\{X_{i,j} > X_{n-\lfloor n(1-\tau_n)\rfloor,n,j}\}\frac{\varphi_{\tau_n}(X_{i,\ell}-\widetilde{\xi}_{\tau_n,\ell})}{(1-\tau_n)\widetilde{\xi}_{\tau_n,\ell}} \\ &- \widehat{\gamma}_{\tau_n,j}\widehat{\gamma}_{\tau_n,\ell}\cdot\frac{1}{n}\sum_{i=1}^n \mathbb{1}\{X_{i,j} > X_{n-\lfloor n(1-\tau_n)\rfloor,n,j}\}\frac{\varphi_{\tau_n}(X_{i,\ell}-\widetilde{\xi}_{\tau_n,\ell})}{(1-\tau_n)\widetilde{\xi}_{\tau_n,\ell}}.\end{aligned}$$

The entry $\boldsymbol{\Sigma}_{j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(2,1)$ is estimated by $\widehat{\boldsymbol{\Sigma}}_{n,j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})(2,1)$ defined in a similar fashion by exchanging $j$ and $\ell$. A proof similar to that of Proposition A.4 shows that the estimator $\widehat{\boldsymbol{\Sigma}}_{n,j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ thus defined is consistent. This suggests an estimator of $\boldsymbol{V}_n^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ defined elementwise as

$$\widehat{\boldsymbol{V}}_{n,j,\ell}^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R}) = \begin{pmatrix} 1 \\ 1/\log d_n \end{pmatrix}^\top \widehat{\boldsymbol{\Sigma}}_{n,j,\ell}^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R}) \begin{pmatrix} 1 \\ 1/\log d_n \end{pmatrix}.$$

We finally deduce an $(1-\alpha)-$asymptotic LAWS-based confidence region for the extreme expectile $\boldsymbol{\xi}_{\tau_n'}$ as the deformed random ellipsoid

$$\widetilde{\mathcal{E}}_{\tau_n',\alpha}^\star = \widetilde{\boldsymbol{\xi}}_{\tau_n'}^\star \exp\left(\frac{\widehat{\boldsymbol{b}}}{\sqrt{n(1-\tau_n)}} + \left[\widehat{\boldsymbol{V}}_n^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})\right]^{1/2} B_d\left(\boldsymbol{0}_d, \sqrt{\chi_{d,1-\alpha}^2/n(1-\tau_n)}\,\log d_n\right)\right).$$

This region is represented on a simulated data set in the bottom panel of Figure I (see Appendix B). A formal result on the coverage level of $\widetilde{\mathcal{E}}_{\tau_n',\alpha}^\star$, which is a consequence of the consistency of $\widehat{\boldsymbol{\Sigma}}_n^{\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$, and hence of $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma},\boldsymbol{R})$ as an estimator of $\boldsymbol{V}^\star(\boldsymbol{\gamma},\boldsymbol{R})$, is the following.

**Theorem 3.4.** *Work under the conditions of Theorem 2.4(i), with additionally $\lambda_j = 0$ for all $j$. If $\boldsymbol{V}^\star(\boldsymbol{\gamma},\boldsymbol{R})$ is positive definite, then $\mathbb{P}(\boldsymbol{\xi}_{\tau_n'} \in \widetilde{\mathcal{E}}_{\tau_n',\alpha}^\star) \to 1-\alpha$ as $n\to\infty$.*

One can easily deduce from that construction a LAWS-based asymptotic $(1-\alpha)-$confidence interval for the $j$th marginal extreme expectile $\xi_{\tau_n',j}$. We assess the finite-sample coverage of this interval in Appendix C.1 in the Supplementary Material document.

**Using the QB extrapolating estimator** We rewrite Equation (7) for $\widehat{\xi}^{\star}_{\tau_n',j}$ as

$$\log\left(\frac{\widehat{\xi}^{\star}_{\tau_n',j}}{\xi_{\tau_n',j}}\right) = (\widehat{\gamma}_{\tau_n,j}-\gamma_j)\log\left(\frac{1-\tau_n}{1-\tau_n'}\right) + \log\left(\frac{(\widehat{\gamma}_{\tau_n,j}^{-1}-1)^{-\widehat{\gamma}_{\tau_n,j}}}{(\gamma_j^{-1}-1)^{-\gamma_j}}\right) + \log\left(\frac{\widehat{q}_{\tau_n,j}}{q_{\tau_n,j}}\right)$$
$$-\log\left(\frac{\xi_{\tau_n',j}}{(\gamma_j^{-1}-1)^{-\gamma_j}q_{\tau_n',j}}\right) - \log\left(\left[\frac{1-\tau_n'}{1-\tau_n}\right]^{\gamma_j}\frac{q_{\tau_n',j}}{q_{\tau_n,j}}\right).$$

By Proposition 1(i) in [9], the first component of the bias on the second line of the right-hand side is essentially a linear combination of $1/q_{\tau_n',j}$ and $A_j((1-\tau_n')^{-1})$. At the extreme level $\tau_n'$, these two terms are typically very small. The second component, meanwhile, is asymptotically proportional to $A_j((1-\tau_n)^{-1})$ [see the proof of Theorem 4.3.8 of 10], and we have discussed previously how estimating this kind of bias component is not necessarily beneficial for confidence region construction. We then ignore these two bias terms and use a Taylor expansion to write, as $n\to\infty$,

$$\log\left(\frac{\widehat{\xi}^{\star}_{\tau_n',j}}{\xi_{\tau_n',j}}\right) \approx (\widehat{\gamma}_{\tau_n,j}-\gamma_j)\left(m(\gamma_j)+\log d_n\right) + \log\left(\frac{\widehat{q}_{\tau_n,j}}{q_{\tau_n,j}}\right) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{n(1-\tau_n)}}\right).$$

Using Theorem 2.2 suggests the following approximation for the QB extrapolating estimator:

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\log\left(\frac{\widehat{\boldsymbol{\xi}}^{\star}_{\tau_n'}}{\boldsymbol{\xi}_{\tau_n'}}\right) \approx \mathcal{N}_d\left(\mathbf{0}_d, \boldsymbol{V}_n^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R})\right), \tag{12}$$

as $n\to\infty$, where $\boldsymbol{V}_n^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R})$ is defined elementwise as

$$\boldsymbol{V}_{n,j,j}^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R}) = \frac{\gamma_j^2}{(\log d_n)^2}(1+(m(\gamma_j)+\log d_n)^2) \text{ for } j=\ell,$$
$$\boldsymbol{V}_{n,j,\ell}^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R}) = \frac{\gamma_j\gamma_\ell}{(\log d_n)^2}\Bigg(R_{j,\ell}(1,1)(m(\gamma_j)+\log(d_n)-1)(m(\gamma_\ell)+\log(d_n)-1)$$
$$+(m(\gamma_j)+\log d_n)\int_0^1 R_{j,\ell}(u,1)\frac{\mathrm{d}u}{u} + (m(\gamma_\ell)+\log d_n)\int_0^1 R_{j,\ell}(1,u)\frac{\mathrm{d}u}{u}\Bigg) \text{ otherwise.}$$

This matrix is readily estimated with the matrix $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R})$ defined as

$$\widehat{\boldsymbol{V}}_{n,j,j}^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R}) = \frac{\widehat{\gamma}_{\tau_n,j}^2}{(\log d_n)^2}(1+(m(\widehat{\gamma}_{\tau_n,j})+\log d_n)^2) \text{ for } j=\ell,$$
$$\widehat{\boldsymbol{V}}_{n,j,\ell}^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R}) = \frac{\widehat{\gamma}_{\tau_n,j}\widehat{\gamma}_{\tau_n,\ell}}{(\log d_n)^2}\Bigg(\widehat{R}_{\tau_n,j,\ell}(1,1)(m(\widehat{\gamma}_{\tau_n,j})+\log(d_n)-1)(m(\widehat{\gamma}_{\tau_n,\ell})+\log(d_n)-1)$$
$$+(m(\widehat{\gamma}_{\tau_n,j})+\log d_n)\int_0^1 \widehat{R}_{\tau_n,j,\ell}(u,1)\frac{\mathrm{d}u}{u} + (m(\widehat{\gamma}_{\tau_n,\ell})+\log d_n)\int_0^1 \widehat{R}_{\tau_n,j,\ell}(1,u)\frac{\mathrm{d}u}{u}\Bigg) \text{ otherwise.}$$

This yields an $(1-\alpha)-$asymptotic QB confidence region for the extreme expectile $\boldsymbol{\xi}_{\tau_n'}$ as the deformed random ellipsoid

$$\widehat{\mathcal{E}}^{\star}_{\tau_n',\alpha} = \widehat{\boldsymbol{\xi}}^{\star}_{\tau_n'} \exp\left(\left[\widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R})\right]^{1/2} B_d\left(\mathbf{0}_d, \sqrt{\chi^2_{d,1-\alpha}/n(1-\tau_n)}\,\log d_n\right)\right).$$

This region is represented alongside the corresponding LAWS-based confidence region $\widetilde{\mathcal{E}}^{\star}_{\tau_n',\alpha}$ in the bottom panel of Figure I in Appendix B. We now establish the asymptotic coverage level of this confidence region in the next theorem, which follows from the consistency of $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}(\boldsymbol{\gamma},\boldsymbol{R})$ as an estimator of $\boldsymbol{V}^{\star}(\boldsymbol{\gamma},\boldsymbol{R})$, and Theorem 2.4(ii).

**Theorem 3.5.** *Work under the conditions of Theorem 2.4(ii), with additionally $\lambda_j = 0$ for all $j$. If $\boldsymbol{V}^{\star}(\boldsymbol{\gamma},\boldsymbol{R})$ is positive definite, then $\mathbb{P}(\boldsymbol{\xi}_{\tau_n'} \in \widehat{\mathcal{E}}^{\star}_{\tau_n',\alpha}) \to 1-\alpha$ as $n\to\infty$.*

We can also deduce from this confidence region a QB asymptotic $(1-\alpha)-$confidence interval for the $j$th marginal extreme expectile at level $\tau_n'$. We study this marginal interval in Appendix C.1 as well.

## 3.3. Testing the equality of extreme expectiles

Another way of carrying out joint inference about several risk measures is to test their equality. This is relevant to actuarial and financial practice, where risk managers may want to assess the asymptotic dependence between several risk variables, individual stock prices or stock indices, as well as whether certain assets or stocks should be considered riskier than others. We show here how our construction of asymptotic confidence regions can be used to design a test of equality of extreme expectiles.

Consider the system of hypotheses

$$\begin{cases} H_0: & \forall j,\ell \in \{1,\ldots,d\},\ \lim_{\tau\uparrow 1} \xi_{\tau,j}/\xi_{\tau,\ell} = 1,\\ H_1: & \exists j,\ell \in \{1,\ldots,d\} \text{ with } j\neq \ell \text{ such that } \lim_{\tau\uparrow 1} \xi_{\tau,j}/\xi_{\tau,\ell} \neq 1.\end{cases}$$

Lemma A.6 in the Appendix shows that the limit of such ratios of expectiles is always well-defined, so that our system of hypotheses makes sense, and in fact

$$\begin{cases} H_0 \Leftrightarrow \forall j,\ell \in \{1,\ldots,d\},\ \lim_{\tau\uparrow 1} q_{\tau,j}/q_{\tau,\ell} = 1,\\ H_1 \Leftrightarrow \exists j,\ell \in \{1,\ldots,d\} \text{ with } j\neq \ell \text{ such that } \lim_{\tau\uparrow 1} q_{\tau,j}/q_{\tau,\ell} \neq 1.\end{cases}$$

We may thus actually view testing the equality of extreme expectiles as a way of testing whether the marginal distributions of $\boldsymbol{X}$ all have the same right tail, or, in other words, as a way of detecting potential heterogeneity in the marginal extremes of the random vector $\boldsymbol{X}$. To construct a testing procedure for this problem, we note that we have at our disposal jointly asymptotically Gaussian estimators of the $\xi_{\tau_n',j}$. Testing the equality of the $\xi_{\tau_n',j}$, which is one way of assessing whether $H_0$ should be rejected in favour of $H_1$, can thus be viewed as testing the equality of the means of a Gaussian random vector. A simple and powerful solution to this problem is given by a likelihood ratio test, which we briefly recall here. Suppose that $\boldsymbol{Z} = (Z_1,\ldots,Z_d)^\top$ is a $d-$dimensional Gaussian random vector with mean $\boldsymbol{m}$ and a *known*, positive definite covariance matrix $\boldsymbol{V}$. Consider the nested models problem

$$\begin{cases} M_0: & m_1 = \cdots = m_d = m,\\ M_1: & \exists (j,\ell) \text{ with } j\neq\ell \text{ such that } m_j \neq m_\ell.\end{cases}$$

The (log-likelihood ratio) deviance statistic for testing the validity of model $M_0$ is

$$\Lambda = (\boldsymbol{Z} - \widehat{m}\mathbf{1}_d)^\top \boldsymbol{V}^{-1}(\boldsymbol{Z} - \widehat{m}\mathbf{1}_d), \ \text{ with } \ \widehat{m} = \frac{\boldsymbol{Z}^\top \boldsymbol{V}^{-1}\mathbf{1}_d}{\mathbf{1}_d^\top \boldsymbol{V}^{-1}\mathbf{1}_d}.$$

In model $M_0$, the statistic $\Lambda$ has a chi-square distribution with $d-1$ degrees of freedom.

In our case, we can set $\boldsymbol{Z}$ to be the LAWS-based extrapolating estimator $\widetilde{\boldsymbol{\xi}}^{\star}_{\tau_n'}$ or the QB extrapolating estimator $\widehat{\boldsymbol{\xi}}^{\star}_{\tau_n'}$. This leads us to two distinct testing procedures.

**LAWS-based test** Following the discussion of Section 3.2 and in particular (10) and (11), we approximate the distribution of the vector $\boldsymbol{Z} = \boldsymbol{Z}_n = \log \widetilde{\boldsymbol{\xi}}^{\star}_{\tau_n'} + \widehat{\boldsymbol{b}}/\sqrt{n(1-\tau_n)}$ by a Gaussian distribution with mean $\boldsymbol{m} = \boldsymbol{m}_n = \log \boldsymbol{\xi}_{\tau_n'}$ and covariance matrix

$$\widehat{\boldsymbol{V}} = \widehat{\boldsymbol{V}}_n = \frac{\log^2[(1-\tau_n)/(1-\tau_n')]}{n(1-\tau_n)} \widehat{\boldsymbol{V}}_n^{\star,\mathrm{LAWS}}(\boldsymbol{\gamma}, \boldsymbol{R})$$

with the notation of Sections 3.1 and 3.2. We thus compute the test statistic

$$\Lambda = \Lambda_n^{\mathrm{LAWS}} = (\boldsymbol{Z} - \widehat{m}\mathbf{1}_d)^\top \widehat{\boldsymbol{V}}^{-1}(\boldsymbol{Z} - \widehat{m}\mathbf{1}_d), \ \text{ with } \ \widehat{m} = \frac{\boldsymbol{Z}^\top \widehat{\boldsymbol{V}}^{-1}\mathbf{1}_d}{\mathbf{1}_d^\top \widehat{\boldsymbol{V}}^{-1}\mathbf{1}_d}.$$

We finally define a test with asymptotic type I error $\alpha$ by deciding that we reject $H_0$ if and only if $\Lambda_n^{\mathrm{LAWS}} > \chi^2_{d-1,1-\alpha}$, where $\chi^2_{d-1,1-\alpha}$ is the $(1-\alpha)-$quantile of the chi-square distribution with $d-1$ degrees of freedom.

**QB test** Still following Section 3.2 and especially (12), we approximate the distribution of the vector $\boldsymbol{Z} = \boldsymbol{Z}_n = \log \widehat{\boldsymbol{\xi}}^{\star}_{\tau_n'}$ by a Gaussian distribution with mean $\boldsymbol{m} = \boldsymbol{m}_n = \log \boldsymbol{\xi}_{\tau_n'}$ and covariance matrix

$$\widehat{\boldsymbol{V}} = \widehat{\boldsymbol{V}}_n = \frac{\log^2[(1-\tau_n)/(1-\tau_n')]}{n(1-\tau_n)} \widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}(\boldsymbol{\gamma}, \boldsymbol{R})$$

with the notation of Section 3.2. We thus compute the test statistic

$$\Lambda = \Lambda_n^{\mathrm{QB}} = (\boldsymbol{Z} - \widehat{m}\mathbf{1}_d)^\top \widehat{\boldsymbol{V}}^{-1}(\boldsymbol{Z} - \widehat{m}\mathbf{1}_d), \ \text{ with } \ \widehat{m} = \frac{\boldsymbol{Z}^\top \widehat{\boldsymbol{V}}^{-1}\mathbf{1}_d}{\mathbf{1}_d^\top \widehat{\boldsymbol{V}}^{-1}\mathbf{1}_d}.$$

A test with asymptotic type I error $\alpha$ is defined by rejecting $H_0$ if and only if $\Lambda_n^{\mathrm{QB}} > \chi^2_{d-1,1-\alpha}$. We conclude this section with a theoretical result that rigorously establishes the asymptotic size of our LAWS-based and QB tests. Up to technical details, the proof consists in applying Cochran's theorem under $H_0$, and interpreting the test statistics as distances between the vector $\boldsymbol{Z}$ and the line $\mathbb{R}\mathbf{1}_d$ spanned by $\mathbf{1}_d$, which must therefore diverge under $H_1$ when this vector concentrates away from $\mathbb{R}\mathbf{1}_d$.

**Theorem 3.6.** *With the notation of Theorem 2.4, assume that $\boldsymbol{V}^{\star}(\boldsymbol{\gamma}, \boldsymbol{R})$ is positive definite.*

*(i) Work under the conditions of Theorem 2.4(i), with additionally $\lambda_j = 0$ for all $j$. Then*

$$\mathbb{P}(\Lambda_n^{\mathrm{LAWS}} > \chi^2_{d-1,1-\alpha}) \to \alpha \text{ under } H_0, \text{ and } \Lambda_n^{\mathrm{LAWS}} \xrightarrow{\mathbb{P}} +\infty \text{ under } H_1.$$

*(ii) Work under the conditions of Theorem 2.4(ii), with additionally $\lambda_j = 0$ for all $j$. Then*

$$\mathbb{P}(\Lambda_n^{\mathrm{QB}} > \chi^2_{d-1,1-\alpha}) \to \alpha \text{ under } H_0, \text{ and } \Lambda_n^{\mathrm{QB}} \xrightarrow{\mathbb{P}} +\infty \text{ under } H_1.$$

Let us finally mention that, for the sake of simplicity, we chose to present a construction of bias and covariance matrix estimators using the quantities $\widehat{\gamma}_{\tau_n,j}$ and $\widehat{R}_{\tau_n,j,\ell}$ as estimators of $\gamma_j$ and $R_{j,\ell}$. Other choices are of course possible, and in particular one may use a different level $\tau$ than the intermediate level $\tau_n$ for the estimation of the covariance matrices. That may improve finite-sample results, as we shall see below in our simulation experiments.

# 4. Simulation experiments

Here we first study the performance of our joint confidence regions for extreme expectiles. We then investigate the power of the two tests for the equality of extreme expectiles. To save space, additional results on the quality of inference about a single extreme expectile and the performance of our joint confidence regions for intermediate expectiles are deferred to Appendices C.1 and C.2, respectively.

We work with, among others, two families of Archimedean copulae, which we briefly introduce below. Further details can be found in [24]. Let $\varphi : (0,1] \to [0,\infty)$ be a convex and strictly decreasing function with $\varphi(1) = 0$ and $\varphi(t) \uparrow \infty$ as $t \downarrow 0$. The Archimedean copula in dimension $d$ with generator $\varphi$ is the $d$-dimensional distribution function $C$ with uniform marginals defined by

$$C(\boldsymbol{u}) = \varphi^{-1}(\varphi(u_1) + \cdots + \varphi(u_d)), \ \boldsymbol{u} = (u_1, \ldots, u_d) \in [0,1]^d.$$

We consider the Clayton family, defined through the generator $\varphi(u) = \theta^{-1}(u^{-\theta} - 1)$ for $\theta > 0$. Here the components of $\boldsymbol{u}$ become independent for $\theta \to 0$, and completely dependent for $\theta \to \infty$. We also consider the Gumbel family, defined through the generator $\varphi(u) = (-\log(u))^{\vartheta}$ for $\vartheta \geq 1$, with $\vartheta = 1$ representing the case of independent variables and $\vartheta \to \infty$ the case of perfectly dependent variables.

The experiments are based on the below models for $\boldsymbol{X} = (X_1, \ldots, X_d)$.

(i) [Clayton-Fréchet model] Let $\boldsymbol{U}$ follow a Clayton copula with dependence parameter $\theta = 10$. Take $X_j = (-\log(U_j))^{-\gamma}$ with $\gamma = 1/3$. Then $\boldsymbol{X}$ has Fréchet marginal distributions with tail index $1/3$ and a Clayton copula dependence structure.

(ii) [Gaussian-Student model] Let $\boldsymbol{U}$ follow a Gaussian copula. The pairwise correlation matrix $\boldsymbol{\rho}$ is defined elementwise as $\rho_{1,2} = 0.8$ for $d = 2$, ($\rho_{1,2} = 0.8$, $\rho_{1,3} = 0.6$, $\rho_{2,3} = 0.4$) for $d = 3$, ($\rho_{1,2} = 0.8$, $\rho_{1,3} = 0.6$, $\rho_{1,4} = 0.4$, $\rho_{2,3} = 0.5$, $\rho_{2,4} = 0.4$, $\rho_{3,4} = 0.4$) for $d = 4$ and ($\rho_{1,2} = 0.8$, $\rho_{1,3} = 0.6$, $\rho_{1,4} = 0.4$, $\rho_{1,5} = 0.2$, $\rho_{2,3} = 0.5$, $\rho_{2,4} = 0.4$, $\rho_{2,5} = 0.3$, $\rho_{3,4} = 0.6$, $\rho_{3,5} = 0.4$, $\rho_{4,5} = 0.3$) for $d = 5$. Take $X_j = F_\nu^{-1}(U_j)$ where $F_\nu$ is the Student-$t$ distribution function with $\nu = 3$ degrees of freedom. Then $\boldsymbol{X}$ has Student-$t$ marginal distributions with tail index $1/3$ and a Gaussian copula dependence structure.

(iii) [Gumbel-Fréchet model] Let $\boldsymbol{U}$ follow a Gumbel copula with dependence parameter $\vartheta = 3$. Take $X_j = (-\log(U_j))^{-\gamma}$ with $\gamma = 1/3$. Then $\boldsymbol{X}$ has Fréchet marginal distributions with tail index $1/3$ and a Gumbel copula dependence structure.

(iv) [Multivariate Student-$t$ model] Let $\boldsymbol{X}$ follow a zero-mean multivariate Student-$t$ distribution with $\nu = 3$ degrees of freedom and a scale matrix $\boldsymbol{\rho}$ as in the Gaussian-Student model (ii) above.

In these four models, all univariate margins have the same tail index $\gamma = 1/3$. The components of $\boldsymbol{X}$ are asymptotically independent in models (i) and (ii), in the sense that all pairwise tail copulae are identically 0, and asymptotically dependent in models (iii) and (iv). It is important to note, however, that even though models (i) and (ii) are technically cases of tail independence, finite samples can and will show a degree of dependence in the joint empirical tail. In each model we simulate $M = 10{,}000$ samples of size $n = 1{,}000$ and dimension $d \in \{2,3,4,5\}$. We estimate and infer the $d$-dimensional expectile $\boldsymbol{\xi}_{\tau_n'}$, with $\tau_n' = 0.999 = 1 - 1/n$, with an anchor intermediate level $\tau_n = 1 - 1/\sqrt{n} \approx 0.968$.

Following the remark at the end of Section 3.3, our confidence regions and test statistics are calculated exactly as we described, except that in $\widehat{\boldsymbol{b}}$, $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{LAWS}}$ and $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}$ and only there, the Hill estimators $\widehat{\gamma}_{\tau_n,j}$ are replaced by their versions $\widehat{\gamma}_{1-k/n,j}$ with $k \in [6, 250]$. The results will be described as a function of this value of $k$. As we shall see, such an external estimation may allow somewhat better finite-sample performance, and it gives some idea about the sensitivity of our procedures to the quality of estimation of the bias and variance components.

## 4.1. Joint inference about intermediate and extreme expectiles

We infer the $d$-dimensional extreme expectile at level $\tau_n'$ using the LAWS and QB extrapolating expectile point estimators and their associated confidence regions $\widetilde{\mathcal{E}}^{\star}_{\tau_n,\alpha}$ and $\widehat{\mathcal{E}}^{\star}_{\tau_n,\alpha}$, with $\alpha = 0.05$ (95% nominal coverage probability). Monte Carlo approximations of the actual non-coverage probabilities are displayed in the first two rows of Figure 1 and compared to $\alpha$. The LAWS-based confidence region seems to perform well, with very stable coverage probabilities close to the nominal level in Fréchet models, and a clearly identified stability region for values of $k$ around 50 with a coverage probability close to the nominal level in Student models. Results seem to be robust with respect to the dimension. The QB confidence region, meanwhile, performs reasonably well, but overall worse than the LAWS confidence region, especially in the two cases involving the Student-$t$ distribution. On that basis and for this sample size, we would recommend using the LAWS confidence region with $k \in [50, 100]$. Let us point out that in every model we consider, $\max_j \rho_j < \max_j(-\gamma_j)$, which is a favourable case for our approach that ignores bias due to the second-order framework. We shall discuss this point again in Section 6.

As a complement, we consider two extra models identical to (iii), except that the first marginal distribution is assumed to have tail index $\gamma_1 = 1/5$ (model (v)) or $\gamma_1 = 2/3$ (model (vi)), with all other margins having the same tail index $1/3$. Results are displayed in the bottom row of Figure 1. The performance of the LAWS confidence region is excellent in both models, while the QB confidence region performs quite well in model (vi) but substantially worse in model (v). The results in model (vi) illustrate a certain robustness of the LAWS approach with respect to departure from our theoretical assumptions, since the variance of the first marginal distribution is infinite, and therefore the asymptotic normality result on which the LAWS confidence region is built does not apply.

## 4.2. Testing the equality of extreme expectiles

We check the performance of the tests for equality of several extreme expectiles. We keep models (i)-(iv) of Section 4.1, although in each case we allow the tail index $\gamma_1$ of the first margin to vary within the interval $[0.15, 0.65]$. The null hypothesis of equal extreme expectiles is then true if and only if $\gamma_1 = 1/3$. Then we perform the LAWS and QB tests and we compute the proportion of rejections, thus deriving Monte Carlo approximations of the type I error probability and power of the tests.

Table 1 reports the type I errors of the LAWS and QB versions of the test. Our tests tend to have a lower type I error than expected, although the LAWS test seems to perform better than the QB test overall, and results obtained with the LAWS version tend to improve as the dimension increases, approaching the nominal level when $d = 5$. Figure 2 displays the power of both versions of the test when as a function of $\gamma_1 \in [0.15, 0.65]$ and $d = 2, 3, 4, 5$. The power curves reflect the reasonable power of both tests, which increases as $\gamma_1$ gets further away from $1/3$. The rejection rate appears to increase for stronger dependence structures, and the highest rejection rate is indeed obtained with the Gumbel-Fréchet model. In the case of the Fréchet distribution, our testing procedures appears to yield reasonably stable results across $k \in [50, 200]$, as Figure 3 shows in the case $d = 5$. With the models involving Student-$t$ marginals (models (ii) and (iv)) the results deteriorate for increasing values of $k$, although reasonably good performance appears to be achieved with $k = 50$. One could of course use different values of $k$ across margins to try to improve results further. We defer this to subsequent work.

## 5. Real data application: risk analysis of multiple exchange rates

We consider negative weekly log-returns (returns for brevity) of the exchange rates of the Great British Pound (GBP) versus the United States Dollar (USD), the Japanese Yen (JPY), the Canadian Dollar (CAD), the Australian Dollar (AUD) and the Norwegian Krone (NOK), from January 1, 1980 to June 26, 2020[1]. These samples of size $n = 2{,}133$ are plotted in the top panels of Figure III in Appendix C of the supplementary file. They are technically, of course, time series data; in the results of this paper we do not enter into the important but difficult question of handling serial dependence. This is the reason why, as suggested by [6], we chose to consider weekly returns as a way to substantially reduce the amount of dependence present in the exchange rates. Since we deal with a fairly long time span, we have also removed the presence of trend and year-over-year seasonality components from each individual series (using the `ts` and `stl` functions, part of the R package `stats`, see [34]) and we therefore focus on the residual series in this analysis. With this in mind, one should see this real data application as an illustration of what the expectile-based methodology can do, rather than as a definitive analysis of this exchange rate data.

The scatterplots in Figure 4 (see also Figure IV in Appendix C) indicate that there is a fairly strong dependence between the GBP-USD and GBP-JPY exchange rates on the one hand, and the GBP-USD and GBP-CAD exchange rates on the other hand. We also find visible dependence within the (GBP-CAD, GBP-NOK) and (GBP-AUD, GBP-NOK) pairs. Table 2 gives estimated correlations between exchange rates (upper off-diagonal), and pairwise extremal coefficient estimates (lower off-diagonal). Recall that the bivariate extremal coefficient is a tail dependence measure $\omega \in [1, 2]$, equal to the value at $(1, 1)$ of the stable tail dependence function [14], with the lower and upper bounds representing the case of complete dependence and independence [see *e.g.* 4]. For two exchange rates labelled $j$ and $\ell$, say, their extremal coefficient is estimated with $\widehat{\omega}_{n,j,\ell} = 2 - \widehat{R}_{\tau_n,j,\ell}(1, 1)$, where $\widehat{R}_{\tau_n,j,\ell}(1, 1)$ is defined in (9). These suggest that there is a fairly strong dependence in the joint tail of the two-dimensional exchange rate returns (GBP-USD, GBP-CAD) and (GBP-CAD, GBP-AUD), with milder dependence in the other pairs of returns.

The purpose of analysing multiple exchange rate returns simultaneously is that it can be useful in understanding and predicting the risks that nations and companies exposed to the global economy are subjected to. Risk analysis is most often based on Value-at-Risk (VaR) at the 99.9% level [see *e.g.* 11, 13] or on a quantile at level $1 - p_n$ where $p_n$ is not larger than $1/n$. We analyse here the joint tail risk in multiple exchange rate returns through our expectile-based multivariate inferential procedures, at the extreme level $\tau_n' = 1 - p_n = 0.9995312$ with $p_n = 1/n$. Point estimates and the 95% confidence intervals of the tail index for the five series are displayed in the bottom row of Figure III in Appendix C. The tails of the individual series seem moderately heavy; estimates are fairly stable for a series-dependent interval of values of $k$. In addition to tail index estimates, Table 2 reports the expectile point estimates obtained with the extrapolating LAWS estimator $\widetilde{\boldsymbol{\xi}}^{\star}_{\tau_n'}$ and QB estimator $\widehat{\boldsymbol{\xi}}^{\star}_{\tau_n'}$. We have also computed two- and three-dimensional asymptotic 95% confidence regions for all pairs and triplets of exchange rate returns, using the LAWS and QB confidence region estimators $\widetilde{\mathcal{E}}^{\star}_{\tau_n',\alpha}$ and

[1] Available from `https://www.investing.com` and on file with the authors.

$\widehat{\mathcal{E}}^{\star}_{\tau'_n,\alpha}$ exactly as described in Section 3.2 (for the sake of simplicity, we do not use external estimates of the $\gamma_j$ as described at the end of Section 4). Figure 4 displays these estimated regions for the most tail dependent pairs and triplets of exchange rate returns (plots for other pairs and triplets are available in Figure IV, see Appendix C). These devices are an important tool for the quantification of the potential contamination risk due to currency fluctuations, and therefore could be useful for risk managers.

Finally, we complete the analysis by performing our testing procedures to assess the validity of the assumption of equal risk severity among exchange rate returns. We did this applying the two versions of the test as described in Section 3.3. By way of comparison, we carried out an analogue test on the equality of extreme quantiles, which is built on the joint asymptotic normality of the Weissman quantile estimators across marginals. Set $\boldsymbol{Z} = \boldsymbol{Z}_n = \log \widehat{\boldsymbol{q}}^{\star}_{\tau'_n}$ as well as

$$\widehat{\boldsymbol{V}} = \widehat{\boldsymbol{V}}_n = \frac{\log^2[(1-\tau_n)/(1-\tau'_n)]}{n(1-\tau_n)} \times \begin{cases} \widehat{\gamma}^2_{\tau_n,j} & \text{if } j = \ell, \\ \widehat{\gamma}_{\tau_n,j}\widehat{\gamma}_{\tau_n,\ell}\widehat{R}_{\tau_n,j,\ell}(1,1) & \text{if } j < \ell. \end{cases}$$

We reject the null hypothesis $H_0$ as in Section 3.3 with asymptotic type I error $\alpha$ when $\Lambda^{\mathrm{Q}}_n > \chi^2_{d-1,1-\alpha}$, where $\Lambda^{\mathrm{Q}}_n = (\boldsymbol{Z} - \widehat{m}\mathbf{1}_d)^\top \widehat{\boldsymbol{V}}^{-1}(\boldsymbol{Z} - \widehat{m}\mathbf{1}_d)$, with $\widehat{m} = (\boldsymbol{Z}^\top \widehat{\boldsymbol{V}}^{-1}\mathbf{1}_d)/(\mathbf{1}_d^\top \widehat{\boldsymbol{V}}^{-1}\mathbf{1}_d)$. The test statistics are represented in Figure 5 as a function of $k$ where $\tau_n = 1 - k/n$. The LAWS and QB test statistics seem roughly stable on the range $k \in [75, 150]$ (which constitutes reasonable ground for taking $k = 150$ as we do in Table 2 and Figure 4). Meanwhile, the statistic $\Lambda^{\mathrm{Q}}_n$ very quickly exhibits an upwards drift. The QB test statistic appears to point to differences in marginal tail behaviour at the $10\%$ level but not at the $5\%$ level, while the LAWS test statistic is even less conclusive. By contrast, the use of the rather unstable statistic $\Lambda^{\mathrm{Q}}_n$ would lead to a rejection at the $5\%$ level. Potential differences in tail behaviour might be due here to the fact that the GBP-JPY exchange rate return could carry different extreme risk than the other returns, see Table 2; this is certainly not obvious either from marginal tail index confidence intervals or extreme expectile confidence intervals, which strongly overlap across marginals, and the results provided by our expectile-based test statistics do not allow one to conclude that these differences exist either. Our methods act here as a valuable complement to the methodology based on comparing extreme quantiles for the analysis of marginal tail behaviour, by indicating that potential differences between the margins are not as clear-cut as they may seem.

## 6. Discussion and perspectives

We obtain in this paper a method that is both reasonably accurate and computationally fast for the construction of joint inference procedures for multiple extreme expectiles. The present paper is written under the assumption that the data are independent and identically distributed, and we do not enter into the difficult question of handling serial dependence, which is why we consider low-frequency data in our real data application. An important future extension of the results of this paper will be their adaptation to the case of stationary but weakly dependent data. This is a very difficult task. Indeed, our method requires investigating the joint asymptotic behaviour of Hill estimators and intermediate order statistics (or intermediate LAWS expectile estimators) across marginals, and the fundamental argument behind the proofs is an asymptotic representation of the Hill estimator as a sum of independent random variables proven in [33], whose validity is currently restricted to the independent context. Results we showed in yet unpublished work in [29], meanwhile, only focus on the estimation of a single extreme expectile for stationary and weakly dependent data, and therefore cannot provide the joint asymptotic behaviour of Hill estimators and intermediate order statistics (or intermediate LAWS expectile estimators) across marginals that we would need to construct our refined asymptotic confidence regions.

Our technique uses both marginal correction and adjusted estimators for the correlation structure between pairs of extreme expectile estimators. We do not correct for second-order approximation-related bias, and in our simulation study we examine different cases when this source of bias is in some sense negligible, which is a favourable case for the methods we introduce. For the second-order approximation-related bias to dominate instead, one should have $|\rho_j| < \gamma_j$ for any $j$, which for the LAWS estimator implies $|\rho_j| < 1/2$ because one needs $\gamma_j < 1/2$ for asymptotic normality of this estimator to hold. When the second-order parameter is this close to 0, the Pareto tail tends to be an unreliable representation of the tail of the underlying distribution. Such cases are known to be very difficult to handle, because the required asymptotic Gaussian approximation might then be poor in finite samples (although, strictly speaking, a precise quantification of the rate of convergence to the Gaussian limit requires more knowledge of the underlying distribution, such as its third-order properties). Our construction, including the simple bias reduction method we suggest, should thus be viewed as a first step towards accurate inference of intermediate and extreme expectiles, but not as a final answer, which would require to handle those difficult cases when the second-order parameters $\rho_j$ are close to 0.

Moreover, in the current implementation, we do not examine the possibility of a correction using a second-order refinement of Condition A(ii), *i.e.* by exploiting the rate of convergence of the dependence structure at intermediate levels to the tail copula dependence structure. An example of such a refinement is Condition 1 in [30, p. 316]. A difficulty in using such a condition for finite-sample adjustments to asymptotic confidence regions is that, just like for marginal corrections using a second-order regular variation condition, it will require using estimators of second-order quantities, whose rate of convergence is typically slow. This might therefore result in conservative confidence regions. Constructing joint inference methodologies that both use the full extent of second-order conditions *via* the use of second-order estimators and which are accurate is a challenging question worthy of future research.

Our methodology is also restricted to the case of a fixed dimension $d$. The case of a growing $d = d_n \to \infty$ is interesting, not least because investigating this case would allow one to capture the impact of having multiple components whose inference errors can add up. This could ultimately result in a better idea of how coverage, type I error and power evolve as the dimension increases, which currently is not clear from either the theory or simulations. This will require proving a completely new set of theoretical results, that in particular finds the trade-off between increasing dimension and the intermediate or extreme level of interest; as such, this line of research is beyond the scope of this paper.

## Acknowledgements

The authors acknowledge an anonymous Associate Editor and two anonymous reviewers for their very helpful comments that led to a much improved version of this paper. Part of this research was carried out when the authors were visiting each other at Bocconi University and the University of Nottingham where G. Stupfler was previously based. Support from the Nottingham PEF Fund is gratefully acknowledged. S.A. Padoan is supported by the Bocconi Institute for Data Science and Analytics (BIDSA), Italy. G. Stupfler is supported by the French National Research Agency under the grant ANR-19-CE40-0013/ExtremReg project. G. Stupfler also acknowledges support from an AXA Research Fund Award on “Mitigating risk in the wake of the COVID-19 pandemic”.

## Supplementary material

The supplement to this article contains technical lemmas and the proofs of all theoretical results in the main article as well as additional details on the construction of confidence regions and further finite-sample results.

# References


[1] ARTZNER, P., DELBAEN, F., EBER, J. M. and HEATH, D. (1999). Coherent measures of risk. *Mathematical Finance* **9** 203–228.

[2] BELLINI, F. and DI BERNARDINO, E. (2017). Risk management with expectiles. *The European Journal of Finance* **23** 487–506.

[3] BELLINI, F., KLAR, B., MÜLLER, A. and GIANIN, E. R. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics* **54** 41–48.

[4] BERANGER, B. and PADOAN, S. A. (2015). Extreme dependence models. In *Extreme Value Modeling and Risk Analysis: Methods and Applications* (D. Dey and J. Yan, eds.) Chapman and Hall/CRC.

[5] CAI, J. J., DE HAAN, L. and ZHOU, C. (2013). Bias correction in extreme value statistics with index around zero. *Extremes* **16** 173–201.

[6] CAI, J. J., EINMAHL, J. H. J., DE HAAN, L. and ZHOU, C. (2015). Estimation of the marginal expected shortfall: the mean when a related variable is extreme. *Journal of the Royal Statistical Society: Series B* **77** 417–442.

[7] 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** 263–292.

[8] DAOUIA, A., GIRARD, S. and STUPFLER, G. (2019). Extreme M-quantiles as risk measures: From $L^1$ to $L^p$ optimization. *Bernoulli* **25** 264–309.

[9] DAOUIA, A., GIRARD, S. and STUPFLER, G. (2020). Tail expectile process and risk assessment. *Bernoulli* **26** 531–556.

[10] DE HAAN, L. and FERREIRA, A. (2006). *Extreme Value Theory: An Introduction*. Springer-Verlag, New York.

[11] DE HAAN, L., MERCADIER, C. and ZHOU, C. (2016). Adapting extreme value statistics to financial time series: dealing with bias and serial dependence. *Finance and Stochastics* **20** 321–354.

[12] DREES, H. (1998). Optimal rates of convergence for estimates of the extreme value index. *Annals of Statistics* **26** 434–448.

[13] DREES, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **9** 617–657.

[14] DREES, H. and HUANG, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *Journal of Multivariate Analysis* **64** 25–46.

[15] 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.

[16] EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin.

[17] GNEITING, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association* **106** 746–762.

[18] 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** 2632–2652.

[19] GOMES, M. I., PESTANA, D. and CAEIRO, F. (2009). A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Statistics & Probability Letters* **79** 295–303.

[20] HALL, P. and WELSH, A. H. (1985). Adaptive estimates of parameters of regular variation. *Annals of Statistics* **13** 331–341.

[21] HILL, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics* **3** 1163–1174.
[22] HOGA, Y. (2018). Detecting tail risk differences in multivariate time series. *Journal of Time Series Analysis* **39** 665–689.
[23] HURLIN, C., LAURENT, S., QUAEDVLIEG, R. and SMEEKES, S. (2017). Risk measure inference. *Journal of Business & Economic Statistics* **35** 499–512.
[24] JOE, H. (2014). *Dependence Modeling with Copulas*. Chapman & Hall/CRC Press, Boca Raton, USA.
[25] KNIGHT, K. (1999). Epi-convergence in distribution and stochastic equi-semicontinuity. Technical Report, University of Toronto.
[26] KOENKER, R. and BASSETT, G. (1978). Regression quantiles. *Econometrica* **46** 33–50.
[27] MAINIK, G., MITOV, G. and RÜSCHENDORF, L. (2015). Portfolio optimization for heavy-tailed assets: Extreme Risk Index vs. Markowitz. *Journal of Empirical Finance* **32** 115–134.
[28] NEWEY, W. K. and POWELL, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica* **55** 819–847.
[29] PADOAN, S. A. and STUPFLER, G. (2020). Extreme expectile estimation for heavy-tailed time series. arXiv:2004.04078v2.
[30] SCHMIDT, R. and STADTMÜLLER, U. (2006). Non-parametric estimation of tail dependence. *Scandinavian Journal of Statistics* **33** 307–335.
[31] SOBOTKA, F. and KNEIB, T. (2012). Geoadditive expectile regression. *Computational Statistics & Data Analysis* **56** 755–767.
[32] STRAETMANS, S. T. M., VERSCHOOR, W. F. C. and WOLFF, C. C. P. (2008). Extreme US stock market fluctuations in the wake of 9/11. *Journal of Applied Econometrics* **23** 17–42.
[33] STUPFLER, G. (2019). On a relationship between randomly and non-randomly thresholded empirical average excesses for heavy tails. *Extremes* **22** 749–769.
[34] R CORE TEAM (2020). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.
[35] WEISSMAN, I. (1978). Estimation of parameters and large quantiles based on the $k$ largest observations. *Journal of the American Statistical Association* **73** 812–815.
[36] ZIEGEL, J. F. (2016). Coherence and elicitability. *Mathematical Finance* **26** 901–918.

**Table 1.** Monte Carlo rejection rate (in %) of the tests of equality of extreme expectiles, with $5\%$ nominal type I error rate, for $n = 1{,}000$, $\tau_n' = 0.999$, under the null hypothesis. The anchor intermediate level is $\tau_n = 1 - 1/\sqrt{n} \approx 0.968$; as described in Section 4, within $\widehat{\boldsymbol{b}}$, $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{LAWS}}$ and $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}$ and only there, the Hill estimators $\widehat{\gamma}_{\tau_n,j}$ are replaced by their versions $\widehat{\gamma}_{1-k/n,j}$, with the third column referring to this value of $k$.

| Model | Method | $k$ | $d=2$ | $d=3$ | $d=4$ | $d=5$ |
|---|---|---|---|---|---|---|
| (i) | LAWS | 50 | 3.13 | 3.45 | 3.78 | 4.22 |
| | QB | 50 | 6.20 | 7.77 | 8.35 | 8.88 |
| (ii) | LAWS | 50 | 2.72 | 3.23 | 3.53 | 3.89 |
| | QB | 50 | 3.56 | 2.24 | 2.08 | 1.90 |
| (iii) | LAWS | 50 | 2.55 | 2.35 | 2.14 | 2.27 |
| | QB | 50 | 3.38 | 3.24 | 2.86 | 3.05 |
| (iv) | LAWS | 50 | 2.26 | 2.73 | 3.78 | 4.09 |
| | QB | 50 | 2.19 | 1.69 | 1.25 | 1.08 |

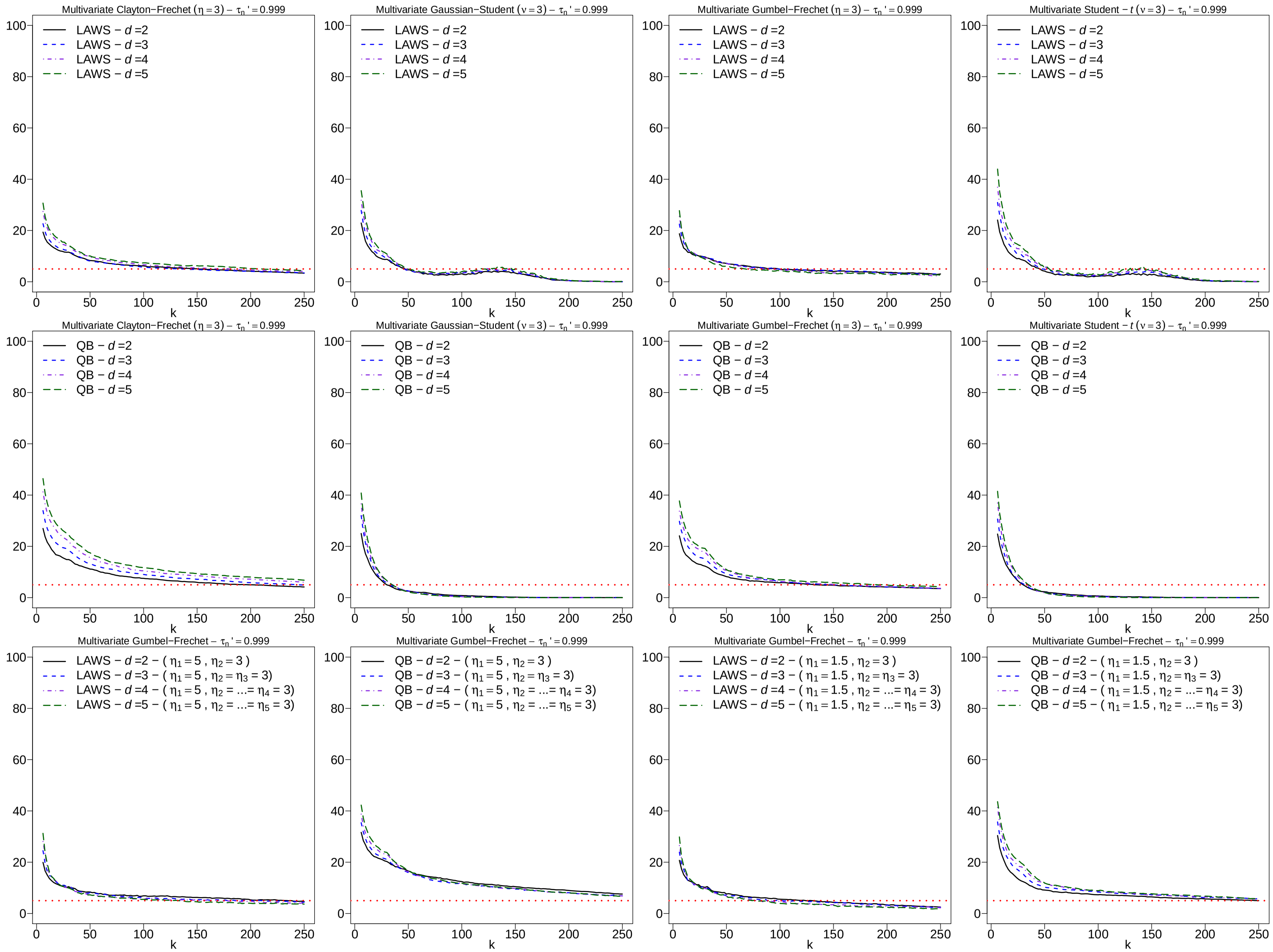


**Figure 1**. Monte Carlo actual non-coverage probabilities (in %) for the LAWS (top row) and QB (middle row) confidence regions $\widetilde{\mathcal{E}}^{\star}_{\tau_n',\alpha}$ and $\widehat{\mathcal{E}}^{\star}_{\tau_n',\alpha}$, with $n = 1{,}000$, $\tau_n' = 0.999$ and 95% nominal level, in models (i)-(iv). The horizontal dotted red line represents the 5% nominal non-coverage probability. In the bottom row, we report the results obtained in model (v) (first plot: LAWS confidence region, second plot: QB confidence region) and (vi) (third plot: LAWS confidence region, fourth plot: QB confidence region). The anchor intermediate level is $\tau_n = 1 - 1/\sqrt{n} \approx 0.968$; as described in Section 4, in $\widehat{\boldsymbol{b}}$, $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{LAWS}}$ and $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}$ and only there, the Hill estimators $\widehat{\gamma}_{\tau_n,j}$ are replaced by their versions $\widehat{\gamma}_{1-k/n,j}$, with the non-coverage probabilities being represented as functions of this value of $k$.

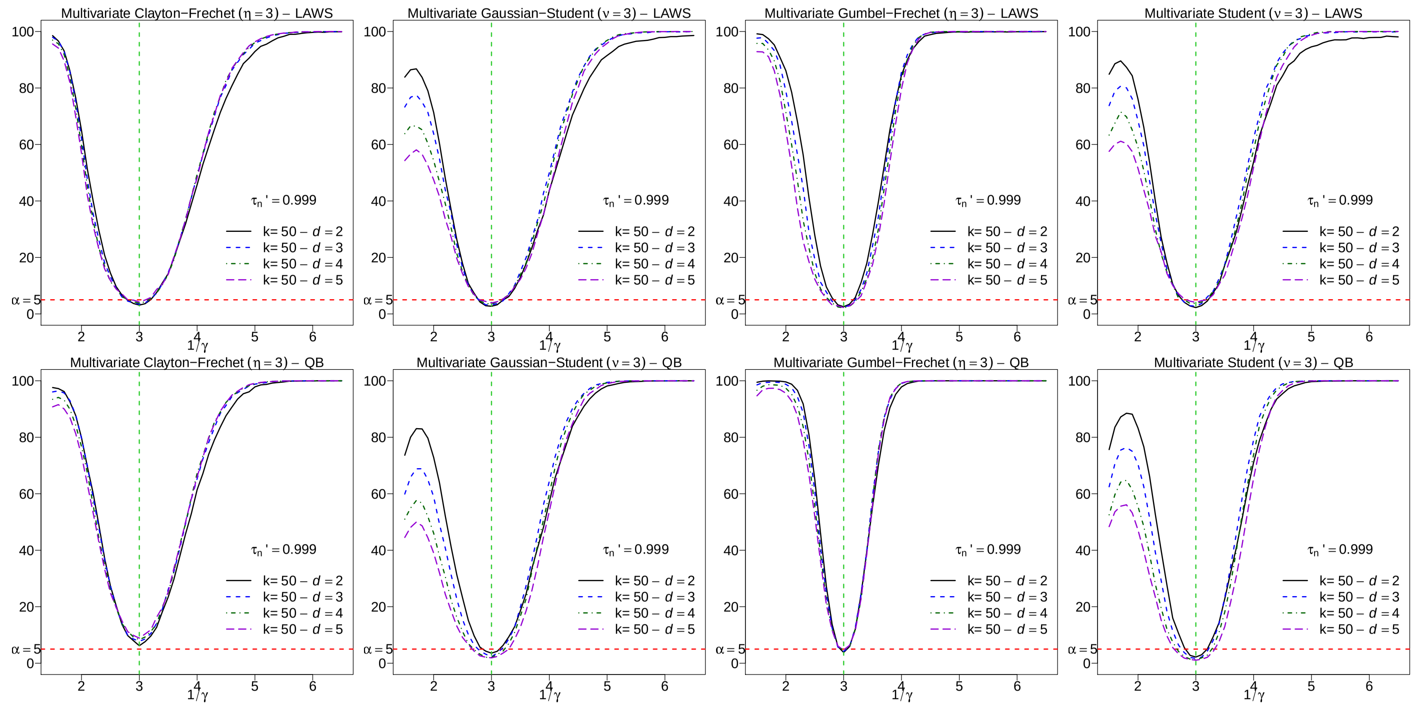


**Figure 2**. Monte Carlo rejection rate (in %) of the LAWS and QB versions of the tests of equality of extreme expectiles, with $5\%$ nominal type I error rate (horizontal dotted red line), for $n = 1{,}000$, $\tau_n' = 0.999$ in models (i)-(iv) where the first marginal distribution is allowed to have a tail index $\gamma_1$ varying in $[0.15, 0.65]$. The anchor intermediate level is $\tau_n = 1 - 1/\sqrt{n} \approx 0.968$; as described in Section 4, in $\widehat{\boldsymbol{b}}$, $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{LAWS}}$ and $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}$ and only there, the Hill estimators $\widehat{\gamma}_{\tau_n,j}$ are replaced by their versions $\widehat{\gamma}_{1-k/n,j}$, here with $k = 50$. The vertical dashed green line represents the value $\gamma = 1/3$ (the case with all marginals having the same tail index).

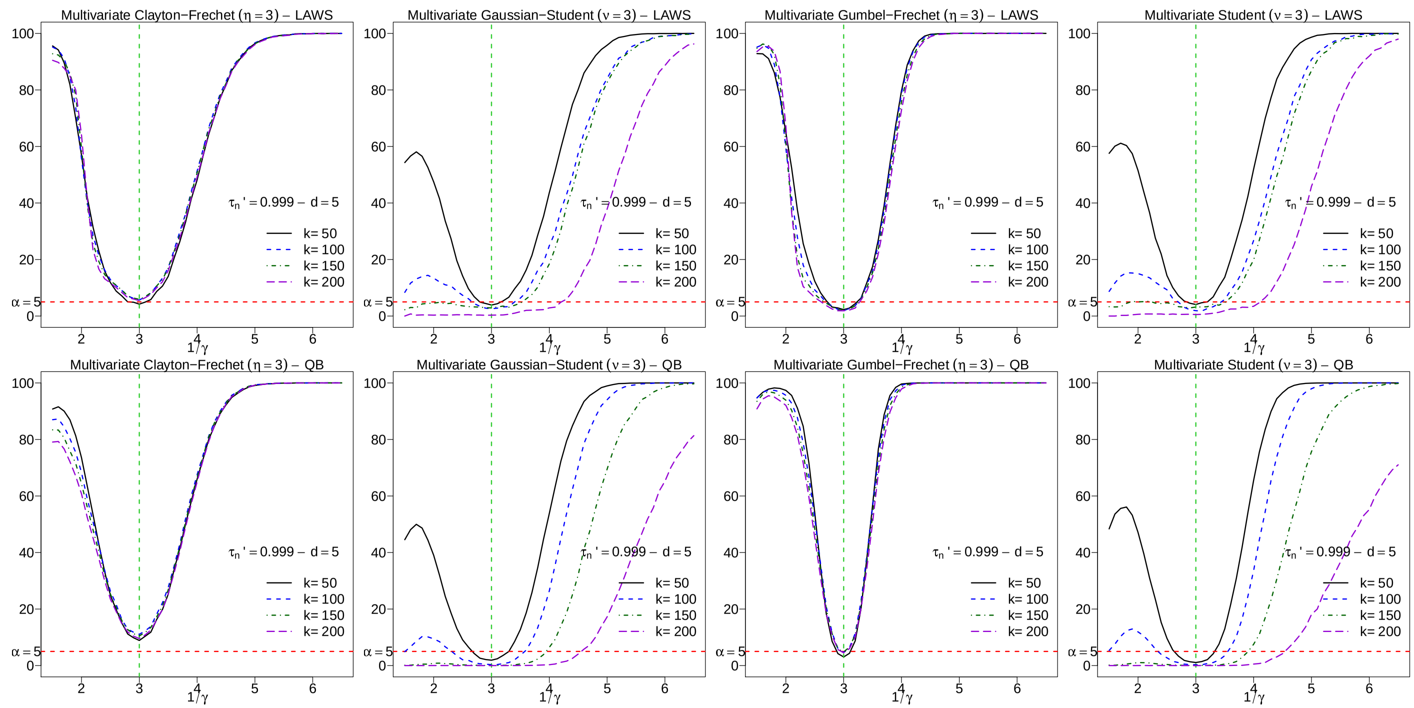


**Figure 3**. Monte Carlo rejection rate (in $\%$) of the LAWS and QB versions of the tests of equality of extreme expectiles, with $5\%$ nominal type I error rate (horizontal dotted red line), for $n = 1{,}000$, $\tau_n' = 0.999$ and dimension $d = 5$, in models (i)-(iv) where the first marginal distribution is allowed to have a tail index $\gamma_1$ varying in $[0.15, 0.65]$. The anchor intermediate level is $\tau_n = 1 - 1/\sqrt{n} \approx 0.968$; as described in Section 4, in $\widehat{\boldsymbol{b}}$, $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{LAWS}}$ and $\widehat{\boldsymbol{V}}_n^{\star,\mathrm{QB}}$ and only there, the Hill estimators $\widehat{\gamma}_{\tau_n,j}$ are replaced by their versions $\widehat{\gamma}_{1-k/n,j}$, with the values of $k = 50, 100, 150$ and $200$ being considered here. The vertical dashed green line represents the value $\gamma = 1/3$ (the case with all marginals having the same tail index).

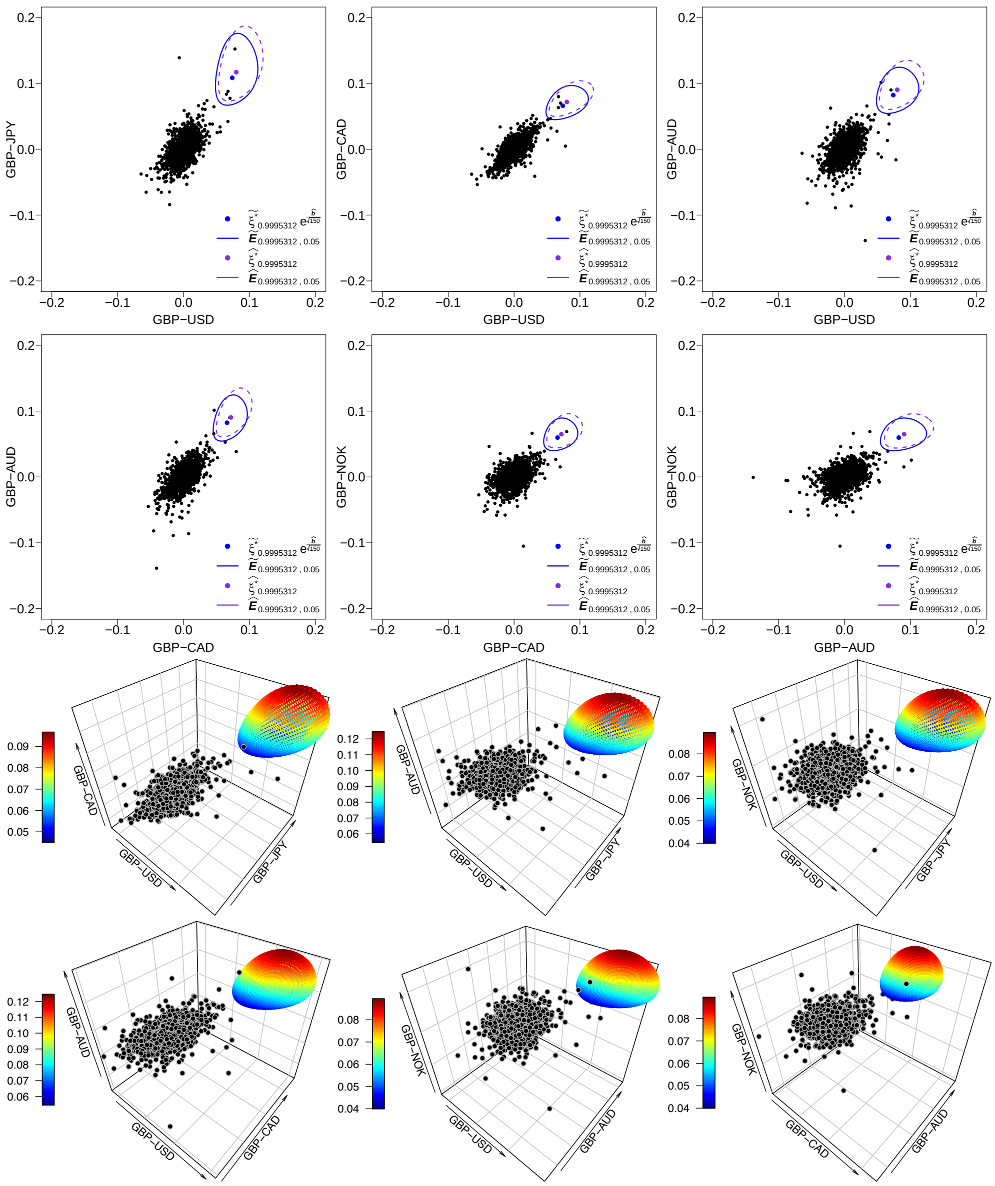


**Figure 4**. Two- and three-dimensional 95% confidence regions estimates for some pairs and triplets of exchange rate returns, obtained with $\tau_n' = 0.9995312$ and $\tau_n = 1 - k/n$ with $k = 150$. In the three-dimensional case we only report LAWS-based confidence regions.

**Table 2.** Top: Pairwise empirical correlations (upper off-diagonal values) and extremal coefficient estimates (lower off-diagonal values) of exchange rate returns. Extremal coefficient estimates are calculated using $\tau_n = 1 - k/n$ where $k = 150$. Bottom: Tail index and extreme expectile estimates relative to the exchange rate returns, obtained with $\tau_n' = 0.9995312$ and $\tau_n = 1 - k/n$ with $k = 150$. Between square brackets are $95\%$ asymptotic confidence intervals. [For $\xi_{\tau_n'}$ these are obtained from the calculation of marginal confidence intervals which is explained in detail at the start of Appendix C.].

| | GBP-USD | GBP-JPY | GBP-CAD | GBP-AUD | GBP-NOK |
|---|---|---|---|---|---|
| GBP-USD | – | 0.561 | 0.753 | 0.488 | 0.346 |
| GBP-JPY | 1.680 | – | 0.385 | 0.272 | 0.367 |
| GBP-CAD | 1.480 | 1.753 | – | 0.648 | 0.442 |
| GBP-AUD | 1.620 | 1.753 | 1.533 | – | 0.406 |
| GBP-NOK | 1.723 | 1.760 | 1.680 | 1.687 | – |

| | | Estimator | |
|---|---|---|---|
| Exchange rate | $\widehat{\gamma}_{1-k/n}$ | $\widetilde{\xi}^{\star}_{\tau_n'}$ | $\widehat{\xi}^{\star}_{\tau_n'}$ |
| GBP-USD | 0.3662 [0.3183, 0.4140] | 0.0711 [0.0484, 0.1044] | 0.0767 [0.0527, 0.1116] |
| GBP-JPY | 0.4422 [0.3845, 0.5000] | 0.1180 [0.0706, 0.1953] | 0.1277 [0.0790, 0.2064] |
| GBP-CAD | 0.3610 [0.3138, 0.4082] | 0.0719 [0.0499, 0.1051] | 0.0788 [0.0548, 0.1134] |
| GBP-AUD | 0.3976 [0.3457, 0.4496] | 0.0926 [0.0631, 0.1378] | 0.1017 [0.0695, 0.1488] |
| GBP-NOK | 0.3787 [0.3292, 0.4281] | 0.0626 [0.0417, 0.0971] | 0.0701 [0.0465, 0.1059] |

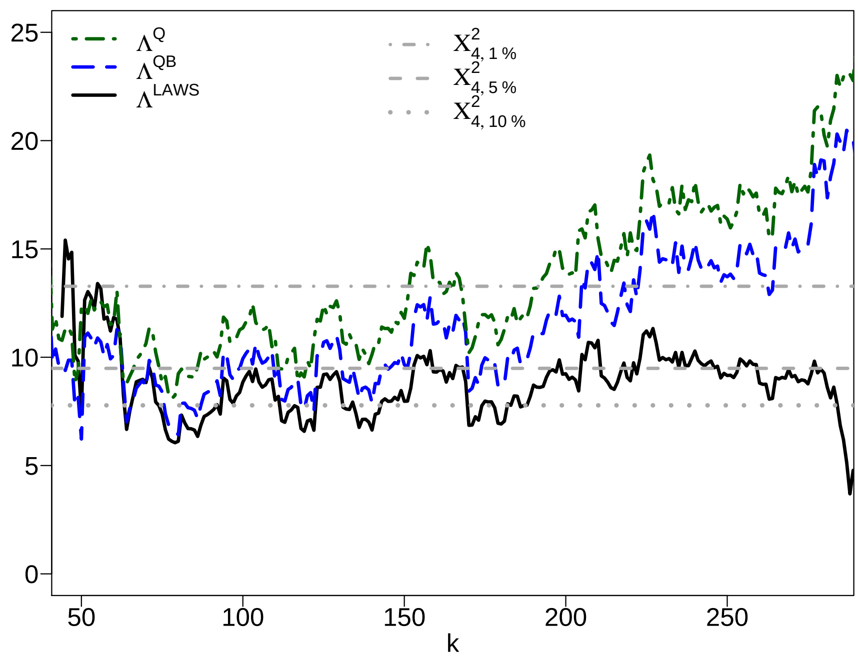


**Figure 5**. Values of the test statistics $\Lambda_n^{\mathrm{LAWS}}$ (solid black line), $\Lambda_n^{\mathrm{QB}}$ (dashed blue line) and $\Lambda_n^{\mathrm{Q}}$ (dashed-dotted green line) on the exchange rates data, as a function of $k \in [50, 250]$, where $\tau_n = 1 - k/n$. The dotted, dashed and dashed-dotted grey lines represent critical values of the tests at the levels $\alpha = 10\%$, $5\%$ and $1\%$ respectively.