WORKING PAPERS

N° 1632

April 2025

“Risk measures beyond quantiles”

Abdelaati Daouia and Gilles Stupfler

Toulouse School of Economics

1

Risk measures beyond quantiles

CONTENTS

The use of quantiles forms the basis of the overwhelming majority of current risk management procedures. Yet, there exist alternative instruments of risk protection that are not (unlike quantiles) based solely on the frequency of tail observations and instead take their severity into account, while adhering to axiomatic requirements. These alternative risk measures have seen increasing interest in the past decade. The current state of development of risk measures beyond quantiles is discussed with a particular focus on three prominent classes: (i) Expected Shortfall (ES) and extremiles, part of the class of spectral and distortion risk measures, (ii) expectiles, which constitute a particular case of generalized M-quantiles, and (iii) systemic risk measures including Marginal Expected Shortfall (MES). A structured overview of their strengths and weaknesses with respect to axiomatic theory, estimation properties, and ease-of-use by risk practitioners will be given. In addition, challenges arising in the asymptotics and mathematical developments will be discussed and the use of each of the ES, extremile, expectile and MES risk measures will be illustrated with real data applications to storm losses in China, tornado losses in the United States, and financial returns series.

1.1 Introduction

Tail risk assessment is concerned with the analysis of rare events that carry potential serious impacts on healthcare systems, the environment or the economy. This includes geohazards and disaster risk, asset/investment risk, systemic risk and emerging risks such as climate, epidemiological, and cybersecurity risks, that are crucial in finance and insurance. The risk of a random position X is usually quantified by a risk measure $M(X)$, where M maps a relevant space of random variables to $\mathbb{R}$. Of practical interest are law-invariant risk measures in the sense that $M(X) = M(Y)$ if the random variables X and Y have the same distribution: all the risk measures we consider throughout this chapter are law-invariant.

In banking and finance, choosing an appropriate risk measure is of great importance. An influential article by [6] provides a meaningful axiomatic foundation for *coherent* risk measures. In this chapter, we adopt the convention that the financial position of interest X is a real-valued random variable, and a positive value of X denotes a loss (*e.g.* X represents negative log-returns). A position Y is then said to be *riskier* than X if $M(Y) \geqslant M(X)$. The risk functional M is said to be coherent if it satisfies the following four requirements:

- Translation equivariance, or equivalently $M(X + a) = M(X) + a$ for all $a \in \mathbb{R}$;
- Positive homogeneity, which amounts to $M(\lambda X) = \lambda M(X)$ for all $\lambda \geqslant 0$;
- Monotonicity, namely, $Y \leqslant X$ with probability 1 implies that $M(Y) \leqslant M(X)$;
- Subadditivity, in the sense that $M(X + Y) \leqslant M(X) + M(Y)$.

An additional important requirement imposed by [1] is *comonotonic additivity*, meaning that $M(X+Y) = M(X)+M(Y)$ for comonotonic random variables X and Y (two random variables X and Y are said to be comonotonic if they can be written as nondecreasing functions of one another). Coherent and comonotonically additive risk measures that are also continuous (see Chapter 4 in [39]) are exactly obtained from *spectral* risk measures of the form

$$M_\phi = M_\phi(X) := \int_0^1 \phi(t)\, q_t \,\mathrm{d}t,$$

where $\phi \in L^1[0, 1]$ is an admissible risk spectrum (*i.e.* non-negative, non-increasing in the L^1-sense, with $\int_0^1 \phi(t)\mathrm{d}t = 1$, see Definitions 2.3 and 2.4 in [1]), and $q_\tau := \inf\{x \in \mathbb{R} : F(x) \geqslant \tau\}$, $\tau \in (0, 1)$, is the τth quantile of X, with F being its distribution function, see Theorems 4.64 p.189 and 4.87 p.200 in [39]. When ϕ is piecewise continuous, [47] show that spectral risk measures belong to the Wang [70] family of *distortion* risk measures

$$M^g = M^g(X) := \int_0^1 q_{1-t}\, \mathrm{d}g(t),$$

where g is a concave, non-decreasing function on $[0, 1]$, with $g(0) = 0$ and $g(1) = 1$. This, in turn, makes it possible to write spectral/distortion risk measures as Choquet integrals, see Definition 4.70 on p.192 in [39].

Arguably, the most common risk measure used in all fields of application is *Value-at-Risk* at level τ (VaR_τ) for $\tau \in (0, 1)$. A main issue with $\mathrm{VaR}_\tau \equiv q_\tau$ in insurance and financial market sectors is its failure to be subadditive in general [1]. It is also often criticized for being unable to account for the size of losses beyond the level τ [17]. A better alternative to VaR_τ in both of these respects is *Expected Shortfall* (ES) at level τ (ES_τ) defined by [3] as

$$\mathrm{ES}_\tau := \frac{1}{1-\tau} \int_\tau^1 q_t \,\mathrm{d}t.$$

When the distribution of X is continuous, this is also known as the τ-Conditional Value-at-Risk, which gives the expectation of X conditional on $X > q_\tau$, namely, $\mathbb{E}(X \mid X > q_\tau)$, see [64]. Being a spectral risk measure generated by the risk aversion function $\phi_\tau(t) = \frac{1}{1-\tau}\mathbb{1}(t \geqslant \tau)$, ES_τ is coherent and comonotonically additive. It is preferred to VaR_τ by practitioners who are concerned with exposure to a catastrophic event, and by major regulators, including the EU, UK, Australia and Canada, which will be requiring the use of $\mathrm{ES}_{97.5\%}$, rather than $\mathrm{VaR}_{99\%}$, in alternative internal models from 1 January 2025. In the EU, this is codified by Article 325ba(1) of Regulation (EU) No 2019/876, which is a revision of the Capital Requirements Regulation (EU) No 575/2013, implementing the latest Basel Committee on Banking Supervision rules.

The ES was criticized though for its conservatism and non-robustness due to its dependency only on the tail event [16, 53], although the debate is very much open about the relevance of robustness in the context of risk measurement [54, 55, 56]. An alternative measure which may steer an advantageous middle course between VaR_τ and ES_τ in terms of sensitivity to the magnitude of extremes is the τth *expectile*

$$e_\tau := \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} \, \mathbb{E}\left\{|\tau - \mathbb{1}(X \leqslant \theta)|(X-\theta)^2 - |\tau - \mathbb{1}(X \leqslant 0)|X^2\right\}$$

which is well-defined, finite and unique as soon as $\mathbb{E}|X| < \infty$. This concept was first introduced by [61] and has been considered as a risk measure by [58]. Being an L^2-minimizer formulated in a way that is analogous to the L^1-formulation of the τth quantile [52],

$$q_\tau \in \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} \, \mathbb{E}\left\{|\tau - \mathbb{1}(X \leqslant \theta)||X-\theta| - |\tau - \mathbb{1}(X \leqslant 0)||X|\right\},$$

the expectile is easy to compute [26]. Its link with ES, as suggested by [68], inspired [22] to introduce a coherent expectile-based form of ES as

$$\mathrm{XES}_\tau := \frac{1}{1-\tau} \int_\tau^1 e_t \, \mathrm{d}t$$

obtained by substituting the expectile e_τ in place of the quantile q_τ in the standard form ES_τ. While expectile-based risk measures have a less natural interpretation and are not comonotonically additive [2], their adoption as reasonable alternatives to VaR_τ and ES_τ has recently attracted a lot of interest, see for instance [8, 9, 20, 32, 57, 63, 73].

Both expectiles and quantiles can be seen as *M-quantiles* [11]. M-quantiles are related to M-functionals of location in the same way quantiles are related to the median [51]. The τth *M-quantile* of a random variable X is essentially defined as

$$\theta_\tau^\psi := \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} \, \mathbb{E}\left\{|\tau - \mathbb{1}(X \leqslant \theta)|\psi(X-\theta) - |\tau - \mathbb{1}(X \leqslant 0)|\psi(X)\right\}$$

based on a suitable loss function ψ. The quantiles q_τ are obtained for the absolute loss function $\psi(x) = |x|$, whereas the expectiles e_τ result from using the quadratic loss function $\psi(x) = x^2$. Other loss functions may be considered, such as the class of *Huber loss functions*

$$\psi_c(x) := \frac{x^2}{2c} \mathbb{1}(|x| < c) + \left(|x| - \frac{c}{2}\right) \mathbb{1}(|x| \geqslant c), \quad \text{for } c > 0,$$

or the L^p-*loss functions* $\psi_p(x) := |x|^p$, for $p \geqslant 1$, which interpolate between the absolute value (for $p = 1$) and quadratic (for $p = 2$) loss functions above. Interestingly, the only M-quantiles that are coherent risk measures are the expectiles [9]. Perhaps one of the strongest arguments in favor of the use of expectiles in practice is given by [7, 66, 73] who proved that expectiles above the mean are the only coherent law-invariant measure of risk which is also *elicitable*, namely, they benefit from a straightforward backtesting methodology, see [8, 35, 36, 42]. While quantiles and expectiles are elicitable functionals, the ES and distortion risk measures different from the Value-at-Risk and the mean are not elicitable. However, non-elicitability does not preclude backtesting since spectral risk measures are actually jointly elicitable with quantiles [38]. This bolstered the interest in a novel risk measure which belongs to the class of spectral risk measures, enjoys various intuitive meanings and affords a reasonable alternative to VaR_τ, e_τ and ES_τ in terms of reactivity to heavy tails, namely, the τth *extremile* of X defined by [18] as

$$x_\tau := \int_0^1 J_\tau(t) \, q_t \, \mathrm{d}t = \int_0^1 q_t \, \mathrm{d}K_\tau(t) = \int_0^1 q_{1-t} \, \mathrm{d}g_\tau(t)$$

for the risk spectrum $J_\tau(\cdot) = K'_\tau(\cdot)$ and distortion function

$$g_\tau(t) = \begin{cases} t^{r(1-\tau)} & \text{if} \quad 0 < \tau \leqslant 1/2, \\ 1 - (1-t)^{r(\tau)} & \text{if} \quad 1/2 \leqslant \tau < 1, \end{cases}$$

with $K_\tau(t) = 1 - g_\tau(1-t)$ and $r(\tau) = \log(1/2)/\log(\tau)$. Extremiles are determined by tail expectations rather than tail probabilities. Similarly to expectiles, they define a least squares analog of quantiles since

$$q_\tau \in \underset{\theta \in \mathbb{R}}{\operatorname{argmin}}\, \mathbb{E}\{J_\tau(F(X))(|X-\theta| - |X|)\},$$

with equality if F is increasing, and x_τ follows by substituting squared deviations in place of the absolute deviations:

$$x_\tau = \underset{\theta \in \mathbb{R}}{\operatorname{argmin}}\, \mathbb{E}\{J_\tau(F(X))(|X-\theta|^2 - |X|^2)\}.$$

Interestingly, x_τ is always, for $\tau \uparrow 1$, more conservative than both q_τ and e_τ, while remaining less pessimistic than ES_τ, for any heavy-tailed distribution [18].

The univariate risk measures described above ignore the interconnection between financial institutions. Following [4], it has been observed that widespread failures and losses of financial firms can impose an externality on the rest of the economy causing systemic risk. Statistical and econometric approaches have been proposed to measure such a global risk by [12] for U.S. financial institutions and [37] for European institutions, culminating in the definition of the SRISK measure. A key component of the SRISK measure for a financial firm is its contribution to a global stock market decline that happens once or twice a decade. This can be measured as the firm's *Marginal Expected Shortfall* (MES): if X denotes the loss return on the firm's equity and Y that of the entire market, the MES is equal to $\mathbb{E}(X|Y > t)$ for a high threshold t reflecting a substantial market decline, typically a tail τth quantile of the distribution of Y as a threshold [4, 12, 14, 37]. The use of the τth expectile as an alternative threshold for quantifying the MES has also been explored by [20, 27], where the asymptotic connection between quantile- and expectile-based MES is unraveled and used as a basis for estimating each tail risk measure from the other. Alternative measures of contagion have been proposed in the literature, including the CoVaR [5], that is, the VaR of the financial system conditional on institutions being under stress. However, none of these measures apart from the MES have yet been treated from the point of view of extreme value theory.

This chapter is structured as follows. Section 1.2 further motivates the consideration of different risk measures, serving different purposes, through disaster losses data and financial returns data. Then Sections 1.3, 1.4 and 1.5 review the properties of, as well as estimation results and extreme value inference about, spectral risk measures, M-quantiles, multivariate and systemic risk measures. The special cases of tail ES, extremile, expectile and MES will be illustrated in each section using the data in Section 1.2. The proposed approaches are fully operational and easy to implement. The datasets as well as our implemented methods and code are freely available at `https://github.com/gillesstupfler`. Section 1.6 concludes and discusses open questions related to estimation and inference of these alternative risk measures.

1.2 Empirical motivation

The China weather losses data consists of Total Damages Adjusted Cost, in billion USD, provided by the EM-DAT[1] international disaster database. The dataset comprises $n = 166$ records of storms, tornadoes, tropical cyclones and hailstorms. The corresponding sample mean, maximum and standard deviation are 817,025,934 USD, 11,447,148,000 USD, and 1,534,618,214 USD, respectively, and the data is highly right-skewed (Figure 1.1(A)). The test of Theorem 5.2.12 in [28] comfortably concludes the heaviness of the upper tail at the three significance levels $0.01, 0.05$ and 0.10 (Figure 1.1(B)). The plot of the bias-reduced Hill estimator of [13] points towards an estimate of the tail index around 0.71 (Figure 1.1(C)), which suggests a very heavy upper tail with an infinite variance. We also consider US tornado monetary losses, in billion USD, provided by the NOAA's Severe Weather Database[2]. We focus on the re-insurance perspective of the loss amounts in excess of 15 million USD, which results in a sample of size $n = 243$ with mean, maximum and standard deviation being 133,760,934 USD, 2,800,100,000 USD, and 330,588,800 USD, respectively. The data is again right-skewed (Figure 1.1(D)) and heavy-tailed (Figure 1.1(E)), with a bias-reduced Hill estimate of the tail index around 0.57 (Figure 1.1(F)).

Besides, our financial data consists of, first, weekly loss returns (minus log-returns) on the equity price of American International Group (AIG) from July 3rd, 2000, to June 30th, 2010. During the 2007-2008 financial crisis, the US government bailed out AIG to avoid jeopardizing the financial integrity of its trading partners, including Goldman Sachs, Morgan Stanley and T. Rowe Price: compared with the weekly loss returns of the latter three financial institutions over the same time period, whose tail indices were found to be less than 1/2 in [14], AIG's returns appear to have a heavier tail, with a bias-reduced Hill estimate around 0.56 (Figure 1.1(G)-(I)). We also consider the weekly loss returns of a value-weighted market index, extracted from [14] over the same time period, by aggregating three US market indices. The tail heaviness of this dataset was checked empirically in [20] with a Hill estimate around 0.37, which is in line with our bias-reduced Hill estimate around 0.39; as a consequence, for this dataset and for the sake of simplicity, we do not formally carry out a series of checks analogous to those of Figure 1.1.

For the four motivating datasets described above, Figure 1.2 displays the empirical versions of the quantile q_τ, the expectile e_τ, the extremile x_τ, and the two ES forms ES_τ and XES_τ, against the security level $\tau \geqslant 0.90$. First, it may be seen that both spectral risk measures, that is, the sample ES (blue) and extremile (red) in this order, provide globally more conservative risk appraisal compared to their M-quantile competitors, namely the sample quantile (orange) and expectile (green). In particular, the sample expectile somewhat surprisingly appears to be the smallest when comparing the sample risk measures at the same level τ. By contrast, the expectile-based form of ES (cyan) is overall much more reactive to the magnitude of extremes than the expectile. Of course, a meaningful comparison requires in practice the use of a different asymmetry level τ for each risk measure. Second, and perhaps most importantly, the sample versions of the quantile-based risk measures q_τ and ES_τ are the only ones to be piecewise constant functions of τ. This awkwardly results in the same risk measurement for substantially different security levels τ. In this respect, the other three competing risk measures based on expectiles and extremiles enjoy a smooth evolution as functions of τ. Finally, and importantly, none of the pure empirical risk measures we examine here are capable of extrapolating beyond the range of the data, when $\tau \geqslant 1 - 1/n$. As such, having to resort to extreme value theory for a finer risk assessment is inevitable.

[1] EM-DAT (Emergency Events Database), CRED / UCLouvain, 2023, Brussels, Belgium – `www.emdat.be`
[2] NOAA (National Oceanic and Atmospheric Administration) – `https://www.spc.noaa.gov/wcm/\#data`

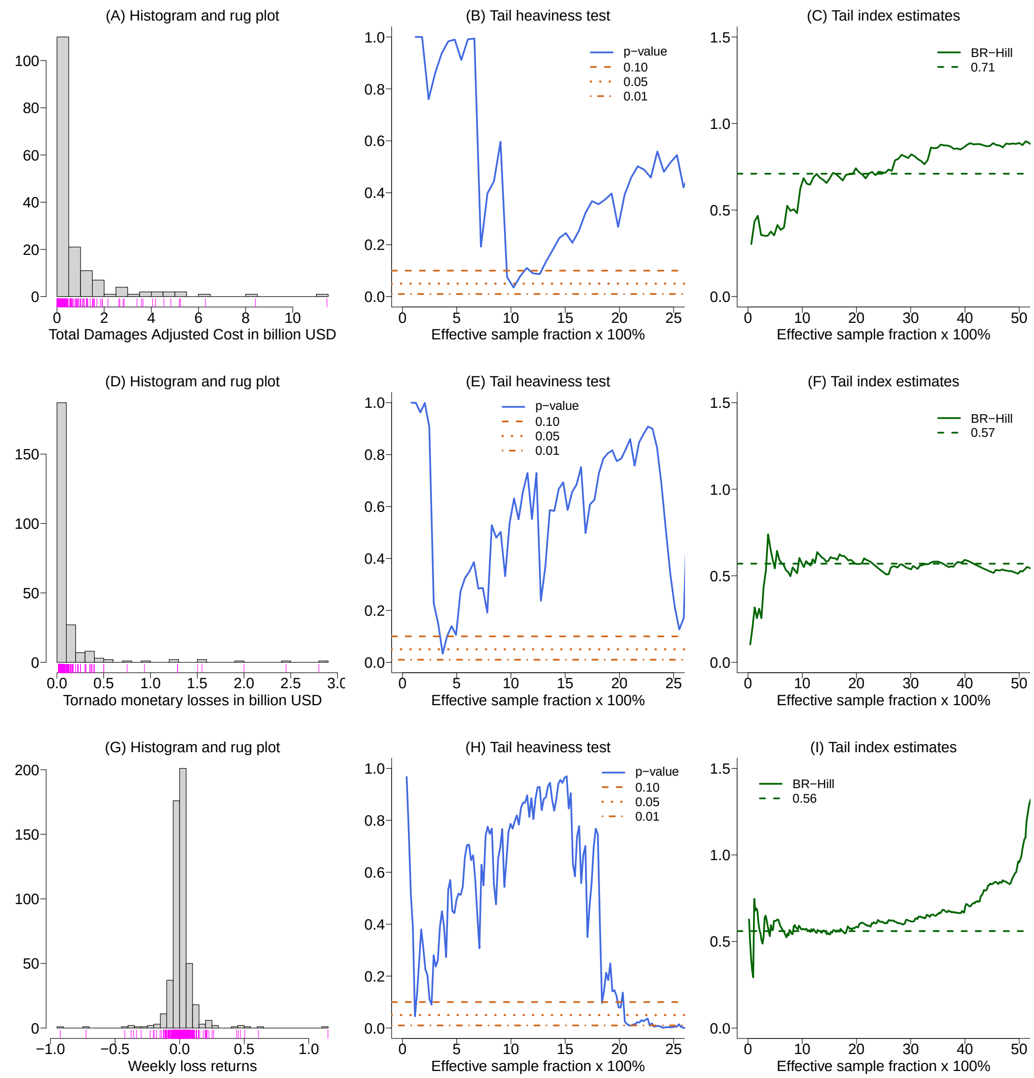

FIGURE 1.1
For the China storm losses data (top panels), US tornado losses data (middle panels), and AIG financial returns data (bottom panels), from left to right: histogram and rug plot of the data, plot of p-values for the tail heaviness test (blue line) along with the three significance levels $0.01, 0.05$ and 0.10 in horizontal lines, and plot of the bias-reduced Hill estimator in solid line, with a first stable region indicated in dashed line. The middle and right plots are graphed as functions of the sample fraction k/n, where k represents the effective sample size of top observations needed for extreme value estimation.

1.3 Spectral risk measures: From Expected Shortfall to extremiles

According to [1], if ϕ is a non-negative, non-increasing function on $[0,1]$ which integrates to 1, the spectral risk measure of X with risk spectrum ϕ is defined as

$$M_\phi = M_\phi(X) = \int_0^1 \phi(t)\, q_t \,\mathrm{d}t.$$

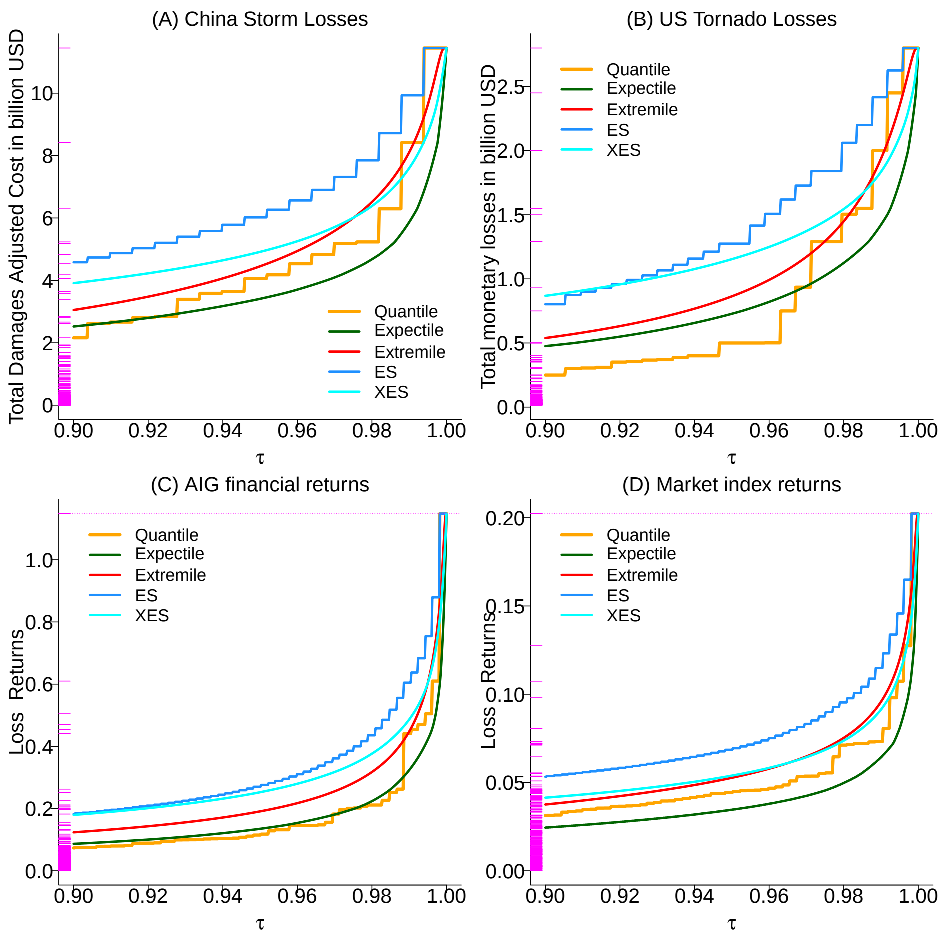

FIGURE 1.2
The sample versions of the quantile q_τ, the expectile e_τ, the extremile x_τ, and the ES forms ES_τ and XES_τ, as functions of $\tau \geqslant 0.9$, for the China storm losses data (A), the US tornado losses data (B), AIG loss returns (C) and the aggregated market index loss returns (D), along with a rug plot of the data on the $y-$axis (for more clarity we only display the positive part of the latter two's rug plots).

Sufficiently regular spectra induce distortion risk measures: when ϕ is the derivative of a smooth function, clearly

$$M_\phi = \int_0^1 q_{1-t}\,\mathrm{d}g(t), \text{ with } \phi(t) = g'(1-t).$$

The motivation for the introduction of distortion risk measures was the calculation of insurance premiums [70]. Important examples include ES as well as the Dual Power and Proportional Hazard risk measures; the Dual Power and Proportional Hazard measures are respectively called MINVAR and MAXVAR by [15] in the context of studying *acceptability indices*. Unlike in the standard setting, the motivation for the use of spectral risk measures in extreme value analysis is not pricing insurance contracts (as it would be overly pessimistic and thus unwise to calculate premiums solely on the basis of infrequent, catastrophic events) but rather the degree of freedom they offer in weighting the most extreme observations, with a view on gathering specific information about disaster risk reflected by the extreme value behavior of a risk variable.

General extreme value adaptations of spectral risk measures The recent literature has considered two extreme value constructions of spectral/distortion risk measures:

1. The excess-of-loss construction in [69]: a pre-specified spectral/distortion risk measure is applied to the distribution of $\max(X - R, 0)$, where R is a high retention level;

2. The conditional construction in [33, 34]: let g be a distortion function and consider, for every $\tau \in (0, 1)$, the distortion risk measure induced by the function g_τ defined as

$$\forall t \in [0,1],\ g_\tau(t) := g\left(\min\left(1, \frac{t}{1-\tau}\right)\right) = \begin{cases} g\left(\dfrac{t}{1-\tau}\right) & \text{if } t \leqslant 1-\tau, \\ 1 & \text{otherwise.} \end{cases}$$

   Considering the distortion risk measure of X with distortion function g_τ then essentially corresponds to applying standard spectral/distortion risk measures to the distribution of X above its τth quantile, that is, the distribution of $X | \{X > q_\tau\}$.

The estimation of these extreme value versions has been considered under the assumption that the data is generated from independent copies of a random variable X having a heavy right tail. More precisely, we assume that the tail quantile function $U : t \mapsto q_{1-1/t}$ of X satisfies the following standard second-order regular variation condition:

$\mathcal{C}_2(\xi, \rho, A)$ For all $x > 0$,

$$\lim_{t\to\infty} \frac{1}{A(t)} \left(\frac{U(tx)}{U(t)} - x^\xi\right) = x^\xi \int_1^x s^{\rho-1} \mathrm{d}s$$

where $\xi > 0$, $\rho \leqslant 0$ and A has constant sign and converges to 0 at infinity. Equivalently

$$\forall x > 0,\ \lim_{t\to\infty} \frac{1}{A(1/\overline{F}(t))} \left(\frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-1/\xi}\right) = \frac{x^{-1/\xi}}{\xi^2} \int_1^x s^{\rho/\xi-1} \mathrm{d}s$$

by Theorem 2.3.9 p.48 in [28], where $\overline{F} = 1 - F$. The function A, such that $|A|$ is regularly varying with index ρ, determines how close the extremes of X are to pure Pareto extremes: the further away ρ is from 0, the closer the extremes of X typically are to pure Pareto extremes, and the easier the extreme value estimation problem is. Any standard heavy-tailed model (Student, Cauchy, Fréchet...) satisfies assumption $\mathcal{C}_2(\xi, \rho, A)$.

Because asymptotic results about the estimation of extremal spectral/distortion risk measures tend to be fairly technical (general results are, for instance, Theorem 3 in [69] and Theorems 1, 2 and 3 in [33]), we specialize our discussion to the ES, which is arguably the most important risk measure covered by both of the constructions we have highlighted above. In the heavy-tailed setting, it is well-known that if $\xi < 1/s$ then $\mathbb{E}(X^s \mathbb{1}(X > 0)) < \infty$, see Exercise 1.16 p.35 in [28]. As a consequence, the ES risk measure is well-defined as soon as $\xi < 1$. Assume here and throughout that the available data points are independent copies $X_1, \ldots, X_n$ of a random variable X satisfying assumption $\mathcal{C}_2(\xi, \rho, A)$. Let $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ be the associated order statistics. Then an obvious estimator of ES_τ is its empirical counterpart

$$\widehat{\mathrm{ES}}_\tau \equiv \widehat{\mathrm{ES}}_{\tau,n} = \frac{1}{1-\tau} \int_\tau^1 \widehat{q}_t \, \mathrm{d}t \quad \text{with} \quad \widehat{q}_t \equiv \widehat{q}_{t,n} = X_{\lceil nt \rceil:n},$$

where $\lceil \cdot \rceil$ is the ceiling function; we drop the subscript n in our estimators throughout for the sake of notational convenience. This estimator is visualized in Figure 1.2 for our motivating datasets at level $\tau \geqslant 0.9$. Since our interest here is in the ES above an extreme level, we take $\tau = \tau_n \uparrow 1$ as $n \to \infty$. A different estimation procedure is suggested by the convergence

$$\frac{\mathrm{ES}_\tau}{q_\tau} \to \frac{1}{1-\xi} \quad \text{as} \quad \tau \uparrow 1, \tag{1.1}$$

(see *e.g.* [69], p.444) which can be used to construct an estimator of ES_{τ_n} provided one has access to an estimator of ξ. A standard choice for that is the Hill estimator [48]: letting $\lfloor \cdot \rfloor$ denote the floor function,

$$\widehat{\xi}_{\tau_n} = \frac{1}{\lfloor n(1-\tau_n) \rfloor} \sum_{i=1}^{\lfloor n(1-\tau_n) \rfloor} \log \frac{X_{n-i+1:n}}{X_{\lceil n\tau_n \rceil:n}}.$$

This is the maximum likelihood estimator in pure Pareto models and arguably the most popular estimator of ξ in heavy-tailed models. One then gets the alternative estimator

$$\widetilde{\mathrm{ES}}_{\tau_n} = \frac{\widehat{q}_{\tau_n}}{1-\widehat{\xi}_{\tau_n}}.$$

The estimators $\widehat{\mathrm{ES}}_{\tau_n}$ and $\widetilde{\mathrm{ES}}_{\tau_n}$ are both asymptotically normal: one path towards a proof is to use a weighted Gaussian approximation to the tail empirical quantile process $s \in [0,1] \mapsto \widehat{q}_{1-(1-\tau_n)s}$, see Theorem 2.4.8 p.52 in [28].

Theorem 1.3.1. *Assume $\mathcal{C}_2(\xi, \rho, A)$ holds. Let $\tau_n \uparrow 1$ with $n(1-\tau_n) \to \infty$.*

(i) If $\xi < 1/2$ and $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) = \mathrm{O}(1)$, then

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\mathrm{ES}}_{\tau_n}}{\mathrm{ES}_{\tau_n}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \frac{2\xi^2(1-\xi)}{1-2\xi}\right).$$

(ii) If $\xi < 1$ and $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda \in \mathbb{R}$, then

$$\sqrt{n(1-\tau_n)}\left(\frac{\widetilde{\mathrm{ES}}_{\tau_n}}{\mathrm{ES}_{\tau_n}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(-\frac{\lambda\xi\rho}{(1-\xi)(1-\rho)(1-\xi-\rho)}, \xi^2\left\{1 + \frac{1}{(1-\xi)^2}\right\}\right).$$

While the first result corresponds to Theorem 2 of [33] and to Corollary 5 of [67] in this form, the second result has not been noted before in the literature and can be viewed as a corollary of Theorem 2.4.8 p.52 in [28]. Let us clearly point out that Theorem 1.3.1 is fundamentally different from the earlier results of [10, 29, 59] which are devoted to the estimation of the ES above a fixed quantile q_τ.

The estimators $\widehat{\mathrm{ES}}_{\tau_n}$ and $\widetilde{\mathrm{ES}}_{\tau_n}$ are consistent and asymptotically normal at intermediate levels τ_n, but cannot extrapolate beyond the range of the observations. At extreme levels τ_n', for which $n(1-\tau_n') = \mathrm{O}(1)$, the key is to use once again (1.1) in conjunction with the classical Weissman [71] approximation to obtain

$$\frac{\mathrm{ES}_{\tau_n'}}{\mathrm{ES}_{\tau_n}} \approx \frac{q_{\tau_n'}}{q_{\tau_n}} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\xi}$$

for all n large enough, which in turn suggests the extrapolated estimators

$$\widehat{\mathrm{ES}}^\star_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\xi}_{\tau_n}} \widehat{\mathrm{ES}}_{\tau_n} \quad \text{and} \quad \widetilde{\mathrm{ES}}^\star_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\xi}_{\tau_n}} \widetilde{\mathrm{ES}}_{\tau_n}.$$

We call them Weissman-type estimators, after [71] who introduced a similar estimator

$$\widehat{q}^{\star}_{\tau'_n} = \left(\frac{1-\tau'_n}{1-\tau_n}\right)^{-\widehat{\xi}_{\tau_n}} \widehat{q}_{\tau_n}$$

for extreme quantiles $q_{\tau'_n}$. They inherit the asymptotic normal distribution of $\widehat{\xi}_{\tau_n}$, at a slightly slower rate of convergence, see Theorem 3 in [33] and Corollary 6 in [67].

Theorem 1.3.2. *Assume $\mathcal{C}_2(\xi, \rho, A)$ holds with $\xi < 1$ and $\rho < 0$. Assume further that $\tau_n, \tau'_n \uparrow 1$ with $n(1-\tau_n) \to \infty$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda \in \mathbb{R}$, $(1-\tau'_n)/(1-\tau_n) \to 0$ and $\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau'_n)) \to \infty$. Then*

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau'_n))}\left(\frac{\widetilde{\mathrm{ES}}^{\star}_{\tau'_n}}{\mathrm{ES}_{\tau'_n}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \xi^2\right).$$

If in addition $\xi < 1/2$, then the same convergence holds true for $\widehat{\mathrm{ES}}^{\star}_{\tau'_n}/\mathrm{ES}_{\tau'_n} - 1$.

We now return to our four motivating data examples to illustrate and compare $\widehat{\mathrm{ES}}^{\star}_{\tau'_n}$ and $\widetilde{\mathrm{ES}}^{\star}_{\tau'_n}$ with the Weissman estimator $\widehat{q}^{\star}_{\tau'_n}$ of $\mathrm{VaR}_{\tau'_n}$ as a benchmark with the extreme level $\tau'_n = 1 - 1/n$. For each dataset, we superimpose in Figure 1.3 the plots of $\widehat{\mathrm{ES}}^{\star}_{\tau'_n}$, $\widetilde{\mathrm{ES}}^{\star}_{\tau'_n}$ and $\widehat{q}^{\star}_{\tau'_n}$ as functions of the effective sample fraction $1-\tau_n$ of the top of the sample used for tail extrapolation in each estimator. The figure also displays the 95% asymptotic confidence intervals derived from the asymptotic normality of $\widetilde{\mathrm{ES}}^{\star}_{\tau'_n}$ in (A)-(D) and of $\widehat{\mathrm{ES}}^{\star}_{\tau'_n}$ only in (D) where $\xi < 1/2$, when ignoring the asymptotic bias by taking $\lambda = 0$. We eyeball the effective sample fraction threshold for stability of the estimates and take $1-\tau_n = 11\%$ for China storm losses, $1-\tau_n = 11\%$ for US tornado losses, $1-\tau_n = 16\%$ for AIG loss returns, and $1-\tau_n = 8\%$ for the market index loss returns, leading to the final pointwise estimates $\widehat{q}^{\star}_{\tau'_n}$, $\widehat{\mathrm{ES}}^{\star}_{\tau'_n}$ and $\widetilde{\mathrm{ES}}^{\star}_{\tau'_n}$ reported in Table 1.1, along with 95% asymptotic confidence intervals. We shall return to the question of asymptotic Gaussian inference and its accuracy in Section 1.6.

Extremiles as a fresh look upon Dual Power risk measures Yet another way of constructing extreme value spectral risk measures is to fix a parametric family of such measures, whose weight shifts towards the right tail as the parameter value converges to the boundary of the parameter space. This is different from the previous construction, where the focus was rather to consider a fixed risk measure applied to a transformation of X. Considering reparametrized probability weighted moment families with this objective in mind is what gives rise to the concept of extremiles, with the reparametrization contributing to a better understanding of some of the properties hidden by their Dual Power formulation.

The original motivation for introducing extremiles comes from the observation in [18] that the τth quantile q_τ coincides with the median of the transformed distribution function $K_\tau(F)$, where we recall that

$$K_\tau(t) = 1 - g_\tau(1-t) = \begin{cases} 1-(1-t)^{r(1-\tau)} & \text{if} \quad 0 < \tau \leqslant 1/2, \\ t^{r(\tau)} & \text{if} \quad 1/2 \leqslant \tau < 1, \end{cases}$$

with $r(\tau) = \log(1/2)/\log(\tau)$. Meanwhile, the mean of this same transformation induces a coherent and more alert risk measure referred to as the τth extremile x_τ that has the closed form expression

$$x_\tau = \frac{\mathbb{E}\{X\, J_\tau(F(X))\}}{\mathbb{E}\{J_\tau(F(X))\}},$$

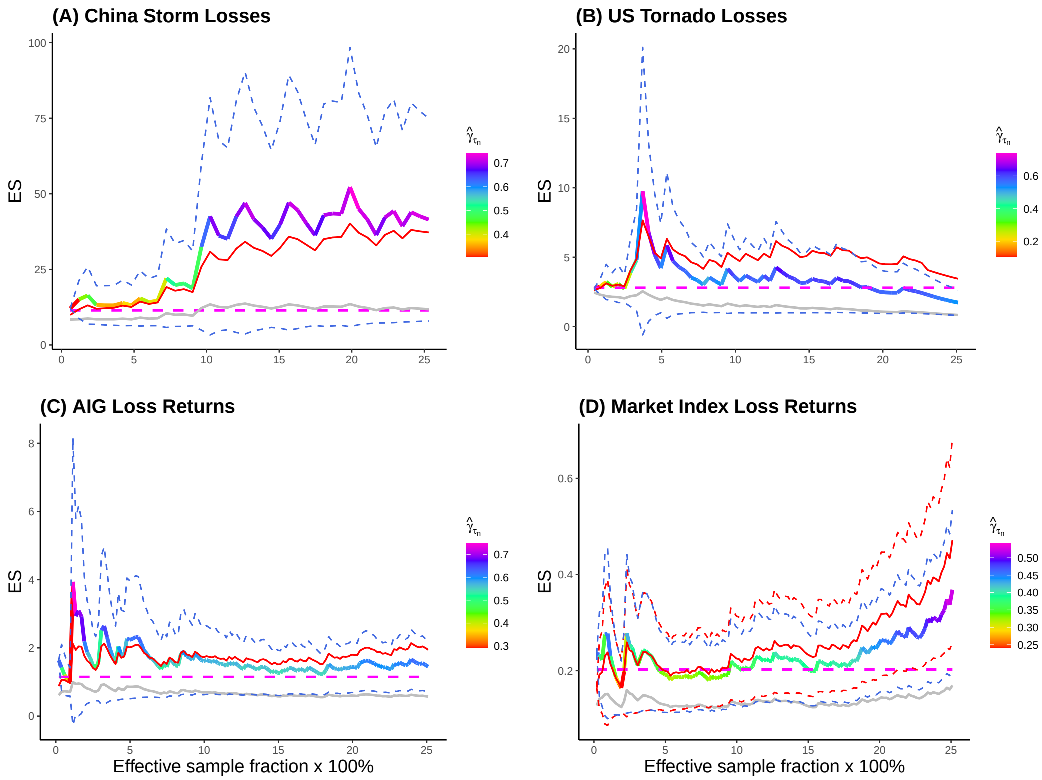

FIGURE 1.3
Extrapolated $\mathrm{ES}_{\tau_n'}$ estimates for (A) China storm losses, (B) US tornado losses, (C) AIG loss returns, and (D) Aggregated US market index loss returns, with $\tau_n' = 1-1/n$: Estimates $\widetilde{\mathrm{ES}}^{\star}_{\tau_n'}$ (rainbow curve, asymptotic 95% confidence intervals in dashed blue), $\widehat{\mathrm{ES}}^{\star}_{\tau_n'}$ (red curve, asymptotic 95% confidence intervals in dashed red) and $\widehat{q}^{\star}_{\tau_n'}$ (gray curve), against the sample fraction $1-\tau_n$. The dashed magenta line is the sample maximum.

which reduces to $x_\tau = \mathbb{E}\{X\, J_\tau(F(X))\}$ for continuous distributions. This risk measure, which is part of the class of Probability-Weighted Moments (PWM) studied in [45, 60], later extended in [30, 31, 46], and discussed in Chapters 5 and 25 of this book, depends on all values of X and puts more weight on extreme realizations, since the weighting density function $J_\tau(\cdot) = K_\tau'(\cdot)$ is increasing for $\tau \geqslant 1/2$ and decreasing for $\tau \leqslant 1/2$. In contrast, q_τ is determined solely by the frequency of tail observations, while ES_τ only depends on the tail event. As visualized in Figure 1.2, the sample extremile, given by the M-estimator

$$\widehat{x}_\tau = \frac{\sum_{i=1}^n J_\tau(i/n) X_{i:n}}{\sum_{i=1}^n J_\tau(i/n)} = \frac{\int_0^1 J_\tau(\lceil nt \rceil/n) X_{\lceil nt \rceil:n}\,\mathrm{d}t}{\frac{1}{n}\sum_{i=1}^n J_\tau(i/n)},$$

exhibits a smooth evolution, steering a middle course between the robustness of ordinary sample quantiles and the severe sensitivity of both empirical ES and extreme quantiles to tail observations. In fact, for Pareto-type distributions with finite first moment, the asymptotic connections (see Propositions 3 and 6 in [18])

$$\frac{x_\tau}{q_\tau} \sim \Gamma(1-\xi)\{\log 2\}^{\xi} > 1 \quad \text{and} \quad \frac{x_\tau}{\mathrm{ES}_\tau} \sim \Gamma(2-\xi)\{\log 2\}^{\xi} < 1 \quad \text{as} \quad \tau \uparrow 1,$$

where Γ is Euler's Gamma function, justify the property, from a risk management viewpoint, that x_τ is always more conservative than q_τ and less pessimistic than ES_τ.

In addition to their duality with quantiles as the mean and the median of the same transformation of X, extremiles also have an intuitive interpretation as expected maxima for $\tau \geqslant 1/2$ and expected minima for $\tau \leqslant 1/2$. More specifically for $\tau \geqslant 1/2$, we have

$$\mathbb{E}\{\max(X_1, \ldots, X_{\lfloor r(\tau) \rfloor})\} \leqslant x_\tau \leqslant \mathbb{E}\{\max(X_1, \ldots, X_{\lceil r(\tau) \rceil})\}$$

where $r(\tau) = \log(1/2)/\log(\tau)$ and $X_1, X_2, \ldots$ are i.i.d. observations from X. In particular, $x_\tau = \mathbb{E}\{\max(X_1, \ldots, X_{r(\tau)})\}$ when $r(\tau) \geqslant 1$ is a integer.

The estimation of tail extremiles at an intermediate level $\tau = \tau_n \uparrow 1$ with $n(1-\tau_n) \to \infty$ as $n \to \infty$, can be done via the M-estimator $\widehat{x}_{\tau_n}$ or by making use of the asymptotic equivalence $x_{\tau_n} \sim \Gamma(1-\xi)\{\log 2\}^{\xi} q_{\tau_n}$, to define the alternative estimator

$$\widetilde{x}_{\tau_n} = \Gamma(1 - \widehat{\xi}_{\tau_n})(\log 2)^{\widehat{\xi}_{\tau_n}} \widehat{q}_{\tau_n}.$$

While the asymptotic normality of $\widehat{x}_{\tau_n}$ has already been established in Theorem 4 of [18], the next theorem is the first result to provide the limit distribution of $\widetilde{x}_{\tau_n}$. It follows as a corollary of Proposition 4 in [18] and Theorem 2.4.8 p.52 in [28]. Let $X_- = \min(X, 0)$.

Theorem 1.3.3. *Assume $\mathcal{C}_2(\xi, \rho, A)$ holds with $\mathbb{E}|X_-| < \infty$. Let $\tau_n \uparrow 1$ with $n(1-\tau_n) \to \infty$.*

(i) If $\xi < 1/2$ and $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) = \mathrm{O}(1)$ then, under the additional regularity assumption that the support of X is an interval, on the interior of which F is twice differentiable with a positive density f such that

$$\sup_{0<t<1} t(1-t)\frac{f'(q_t)}{\{f(q_t)\}^2} < \infty \quad \textit{and} \quad \lim_{t\to+\infty} t\frac{f(t)}{1-F(t)} = \xi,$$

one has

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{x}_{\tau_n}}{x_{\tau_n}} - 1\right) \xrightarrow{\mathrm{d}} \frac{\xi\sqrt{\log 2}}{\Gamma(1-\xi)} \int_0^\infty e^{-s} s^{-\xi-1} B(s)\mathrm{d}s = \xi\sqrt{\log 2}\int_0^\infty g(s|1-\xi)B(s)\frac{\mathrm{d}s}{s}$$

where B is a standard Brownian motion and $g(s|a) = \frac{1}{\Gamma(a)}e^{-s}s^{a-1}$, for $s > 0$, is the density function of the Gamma distribution with unit rate parameter and scale parameter $a > 0$. The limiting distribution is centered and has variance

$$2\xi^2 \log 2 \int_0^\infty G(t|1-\xi)g(t|1-\xi)\frac{\mathrm{d}t}{t} = \xi^2 \log 2 \int_0^\infty \left(\frac{G(t|1-\xi)}{t}\right)^2 \mathrm{d}t$$

where $G(t|a) = \int_0^t g(s|a)\mathrm{d}s = \int_0^t \frac{1}{\Gamma(a)}e^{-s}s^{a-1}\mathrm{d}s$ is the distribution function of the Gamma distribution with unit rate parameter and scale parameter $a > 0$.

(ii) If $\xi < 1$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}(1-\tau_n) \to \nu \in \mathbb{R}$, then

$$\sqrt{n(1-\tau_n)}\left(\frac{\widetilde{x}_{\tau_n}}{x_{\tau_n}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\lambda\left\{\frac{1}{1-\rho}\left(\log(\log 2) - \frac{\Gamma'(1-\xi)}{\Gamma(1-\xi)}\right) - C_1(\xi,\rho)\right\} - C_2(\xi)\nu,\right.$$

$$\xi^2\left\{1+\left(\log(\log 2)-\frac{\Gamma'(1-\xi)}{\Gamma(1-\xi)}\right)^2\right\}\Bigg).$$

Here

$$C_1(\xi,\rho)=\begin{cases}\dfrac{1}{\rho}\left(\dfrac{\Gamma(1-\xi-\rho)(\log 2)^{\xi+\rho}}{\Gamma(1-\xi)(\log 2)^{\xi}}-1\right) & \text{if } \rho<0,\\[2ex] \dfrac{\int_0^\infty e^{-t}t^{-\xi}\left(\log(\log 2)-\log(t)\right)\mathrm{d}t}{\Gamma(1-\xi)} & \text{otherwise,}\end{cases}$$

and $C_2(\xi)=\dfrac{\xi}{2}\left(\dfrac{1-\xi}{\log 2}-1\right)$.

At the far tail, for extreme levels τ_n' with $n(1-\tau_n')=\mathrm{O}(1)$, combining the asymptotic connection between tail extremiles and quantiles with Weissman's [71] approximation gives

$$\frac{x_{\tau_n'}}{x_{\tau_n}}\approx\frac{q_{\tau_n'}}{q_{\tau_n}}\approx\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\xi},$$

which motivates the following extreme value estimators of $x_{\tau_n'}$:

$$\widehat{x}^{\star}_{\tau_n'}=\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\xi}_{\tau_n}}\widehat{x}_{\tau_n}\quad\text{and}\quad\widetilde{x}^{\star}_{\tau_n'}=\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\xi}_{\tau_n}}\widetilde{x}_{\tau_n}.$$

Their convergence is established in Theorems 3 and 5 of [18].

Theorem 1.3.4. *Assume $\mathbb{E}|X_-|<\infty$ and $\mathcal{C}_2(\xi,\rho,A)$ holds with $\xi<1$ and $\rho<0$. Assume further that $\tau_n,\tau_n'\uparrow 1$ with $n(1-\tau_n)\to\infty$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1})\to\lambda\in\mathbb{R}$, $\sqrt{n(1-\tau_n)}(1-\tau_n)\to\nu\in\mathbb{R}$, $(1-\tau_n')/(1-\tau_n)\to 0$ and $\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))\to\infty$. Then*

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\widetilde{x}^{\star}_{\tau_n'}}{x_{\tau_n'}}-1\right)\xrightarrow{\mathrm{d}}\mathcal{N}\left(\frac{\lambda}{1-\rho},\xi^2\right).$$

If $\xi<1/2$ and the regularity conditions of Theorem 1.3.3(i) moreover hold, then the same convergence result holds true for $\widehat{x}^{\star}_{\tau_n'}/x_{\tau_n'}-1$.

Figure 1.4 shows the evolution of these two competing estimates with respect to their quantile analog $\widehat{q}^{\star}_{\tau_n'}$, as functions of the sample fraction $1-\tau_n$, for the extreme level $\tau_n'=1-1/n$, along with the 95% asymptotic confidence intervals derived from the asymptotic normality of $\widetilde{x}^{\star}_{\tau_n'}$ in (A)-(D) and of $\widehat{x}^{\star}_{\tau_n'}$ in (D), assuming that $\lambda=0$. The final pointwise estimates, obtained from the same effective sample fraction threshold (selected and described above) for stability of the benchmark $\widehat{q}^{\star}_{\tau_n'}$, are reported in Table 1.1, along with their asymptotic 95% confidence intervals.

1.4 M-quantile risk measures: From L^1 to convex optimization

The extremile is obtained as the unique minimizer of an expected loss function depending on the unknown distribution function of the observations. This is the key reason why extremiles are not elicitable in the sense of [42]. Of interest, therefore, are risk measures defined as

$$\underset{\theta\in\mathbb{R}}{\operatorname{argmin}}\,\mathbb{E}\left\{W(\tau,X-\theta)|X-\theta|^p\right\}$$

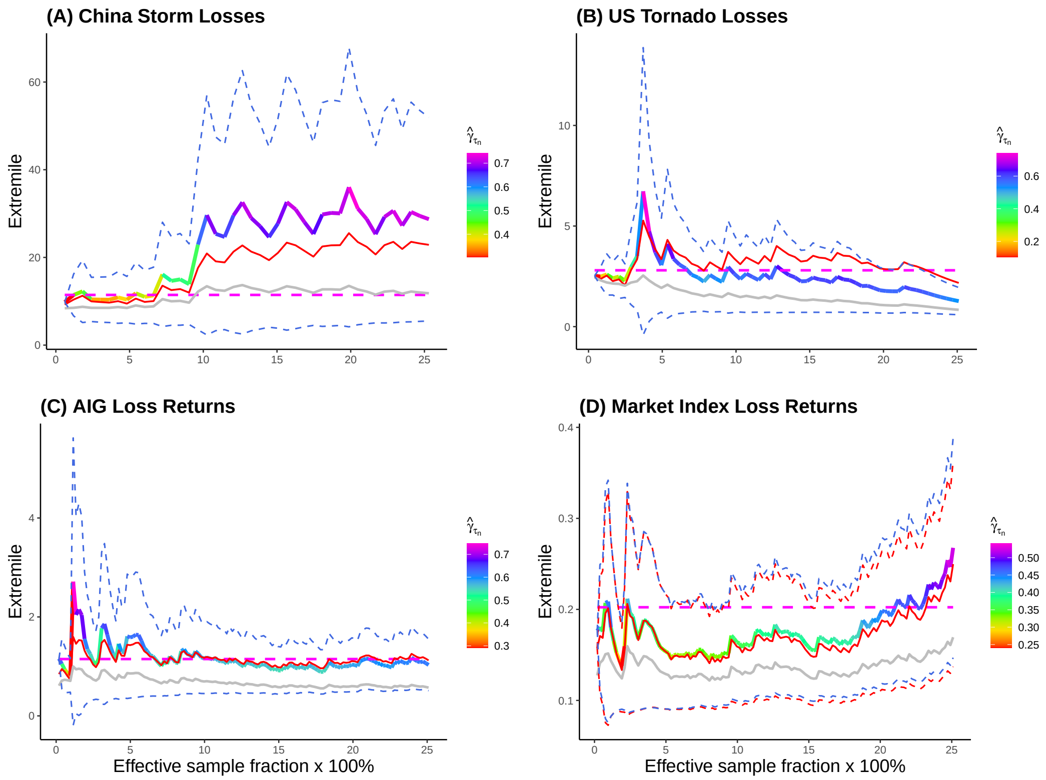

FIGURE 1.4
Extrapolated $x_{\tau_n'}$ estimates for (A) China storm losses, (B) US tornado losses, (C) AIG loss returns, and (D) Aggregated US market index loss returns, with $\tau_n' = 1 - 1/n$: Estimates $\widetilde{x}^{\star}_{\tau_n'}$ (rainbow curve, asymptotic 95% confidence intervals in dashed blue), $\widehat{x}^{\star}_{\tau_n'}$ (red curve, asymptotic 95% confidence intervals in dashed red) and $\widehat{q}^{\star}_{\tau_n'}$ (gray curve), against the sample fraction $1 - \tau_n$. The dashed magenta line is the sample maximum.

where W is a weighting function formalizing the asymmetrical way in which the left and right tails of X are taken into account, and $p \geqslant 1$ encodes how robust the risk measure will be to extreme observations, with increasing p standing for decreasing robustness. Choosing $W(\tau, v) = |\tau - \mathbb{1}(v \leqslant 0)|$ and $p = 1$ produces the usual quantiles [52]. This same choice of W, with varying $p \in [1, \infty)$, generates the class of L^p-quantiles, studied from the extreme value perspective in [21]. These risk measures are law-invariant and elicitable for any $p \geqslant 1$ but, according to [73], only the expectiles, obtained for $p = 2$, induce a coherent risk functional. For this reason, we focus hereafter on the case $p = 2$.

Expectile-based Value-at-Risk In the intermediate case when $\tau = \tau_n \uparrow 1$ and $n(1 - \tau_n) \to \infty$ as $n \to \infty$, a direct, Least Asymmetrically Weighted Squares (LAWS) estimator of the expectile e_{τ_n} is given by its empirical version

$$\widehat{e}_{\tau_n} = \underset{\theta \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^{n} |\tau_n - \mathbb{1}(X_i \leqslant \theta)| (X_i - \theta)^2.$$

Its asymptotic normality is proven in Theorem 1 of [22] for heavy-tailed distributions with $\xi < 1/2$. In the more general setting where $0 < \xi < 1$, an indirect, quantile-based estimator of e_{τ_n} is obtained from the asymptotic connection $e_{\tau_n} \sim (\xi^{-1}-1)^{-\xi} q_{\tau_n}$, $n \to \infty$, as

$$\widetilde{e}_{\tau_n} = (\widehat{\xi}_{\tau_n}^{-1} - 1)^{-\widehat{\xi}_{\tau_n}} \widehat{q}_{\tau_n}.$$

Its asymptotic normality is established in Corollary 2 of [20].

Theorem 1.4.1. *Assume $\mathcal{C}_2(\xi, \rho, A)$ holds and let $\tau_n \uparrow 1$ with $n(1-\tau_n) \to \infty$.*

(i) If $\xi < 1/2$, $\mathbb{E}|X_-|^2 < \infty$ and $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) = \mathrm{O}(1)$, then

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{e}_{\tau_n}}{e_{\tau_n}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \frac{2\xi^3}{1-2\xi}\right).$$

(ii) If $\xi < 1$, $\mathbb{E}|X_-| < \infty$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}/q_{\tau_n} \to \mu \in \mathbb{R}$, then

$$\sqrt{n(1-\tau_n)}\left(\frac{\widetilde{e}_{\tau_n}}{e_{\tau_n}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}(b(\xi,\rho), \xi^2[1 + \{(1-\xi)^{-1} - \log(\xi^{-1}-1)\}^2]),$$

with

$$b(\xi,\rho) = \left(\frac{(1-\xi)^{-1} - \log(\xi^{-1}-1)}{1-\rho} - \frac{(\xi^{-1}-1)^{-\rho}}{1-\xi-\rho} - \frac{(\xi^{-1}-1)^{-\rho}-1}{\rho}\right)\lambda$$
$$- \xi(\xi^{-1}-1)^{\xi}\mathbb{E}(X)\mu.$$

While the empirical version $\widehat{e}_{\tau_n}$ of the risk measure e_{τ_n} inherits its property of coherence and has a smaller asymptotic variance than the quantile-based competitor $\widetilde{e}_{\tau_n}$ for low tail index values $\xi < \xi_0 \approx 0.262$, as can also be seen from Figure 1.10 below, the latter estimator is (asymptotically) more efficient over the range of values $\xi > \xi_0$. In order to obtain the best of these two estimators, [22] suggested to combine them by considering the weighted version $\overline{e}_{\tau_n}(\alpha) = \alpha\,\widehat{e}_{\tau_n} + (1-\alpha)\widetilde{e}_{\tau_n}$, for $\alpha \in \mathbb{R}$, whose asymptotic distribution is established in Theorem 4 therein. The weight α minimizing the asymptotic variance of $\overline{e}_{\tau_n}(\alpha)$ was determined in [23], along with an adaptive estimator for e_{τ_n} obtained from plugging the variance-optimal weight α into the linear combination $\overline{e}_{\tau_n}(\alpha)$. Monte Carlo evidence from [23] suggests that the resulting adaptive estimator indeed performs well.

In the extreme case when $\tau = \tau_n'$ with $n(1-\tau_n') = \mathrm{O}(1)$, combining the asymptotic proportionality relationship between expectiles and quantiles with the Weissman approximation yields

$$\frac{e_{\tau_n'}}{e_{\tau_n}} \sim \frac{q_{\tau_n'}}{q_{\tau_n}} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\xi},$$

which motivates the class of expectile Weissman-type estimators

$$\overline{e}^{\star}_{\tau_n'}(\alpha) = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\xi}_{\tau_n}} \overline{e}_{\tau_n}(\alpha)$$

based on the weighted intermediate estimators. Such extrapolated estimators inherit the asymptotic normal distribution of the tail index estimator $\widehat{\xi}_{\tau_n}$, as shown in Theorem 5 of [22]. Of particular interest are the purely indirect estimator $\widetilde{e}^{\star}_{\tau_n'} := \overline{e}^{\star}_{\tau_n'}(0)$ and direct estimator $\widehat{e}^{\star}_{\tau_n'} := \overline{e}^{\star}_{\tau_n'}(1)$, which correspond to the two special cases $\alpha = 0$ and $\alpha = 1$, and whose asymptotic normality properties were established, respectively, for $\xi < 1$ and $\xi < 1/2$ in Corollary 3 and Corollary 4 of [20].

Theorem 1.4.2. *Assume $\mathcal{C}_2(\xi,\rho,A)$ holds with $\rho<0$, and let $\tau_n,\tau_n'\uparrow 1$ be such that $n(1-\tau_n)\to\infty$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1})\to\lambda\in\mathbb{R}$, $\sqrt{n(1-\tau_n)}/q_{\tau_n}\to\mu\in\mathbb{R}$, $(1-\tau_n')/(1-\tau_n)\to 0$ and $\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))\to\infty$.*

(i) If $\xi<1/2$ and $\mathbb{E}|X_-|^2<\infty$, then for any $\alpha\in\mathbb{R}$,

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\overline{e}^{\star}_{\tau_n'}(\alpha)}{e_{\tau_n'}}-1\right)\xrightarrow{\mathrm{d}}\mathcal{N}\left(\frac{\lambda}{1-\rho},\xi^2\right).$$

(ii) This remains valid for $\alpha=0$ under the weaker assumptions $\xi<1$ and $\mathbb{E}|X_-|<\infty$.

For our motivating datasets, the competing estimates $\widehat{e}^{\star}_{\tau_n'}$ and $\widetilde{e}^{\star}_{\tau_n'}$ are graphed with their quantile analog $\widehat{q}^{\star}_{\tau_n'}$ in Figure 1.5, as functions of the sample fraction $1-\tau_n$, for the extreme level $\tau_n'=1-1/n$, along with the 95% asymptotic confidence intervals associated with $\widetilde{e}^{\star}_{\tau_n'}$ in (A)-(D) and with $\widehat{e}^{\star}_{\tau_n'}$ in (D), under the bias condition $\lambda=0$. The final pointwise estimates $\widehat{e}^{\star}_{\tau_n'}$ and $\widetilde{e}^{\star}_{\tau_n'}$, selected by using the same effective sample fraction threshold for stability of $\widehat{q}^{\star}_{\tau_n'}$, are reported in Table 1.1 with their associated confidence intervals.

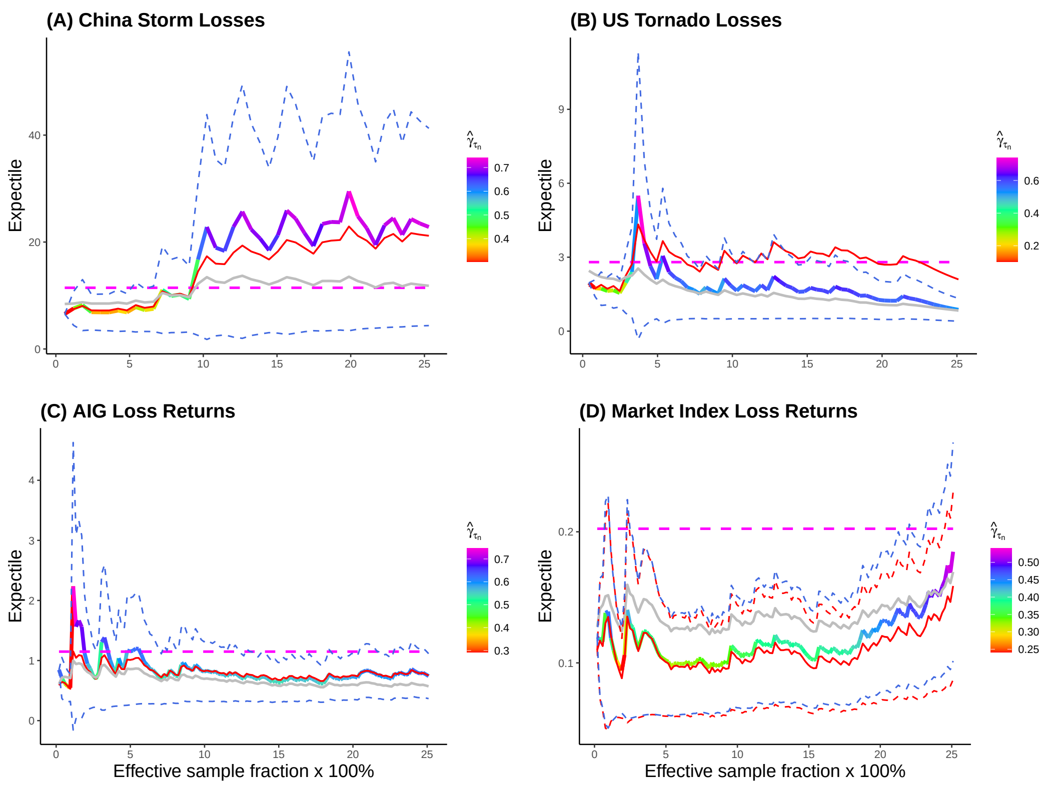

FIGURE 1.5
Extrapolated $e_{\tau_n'}$ estimates for (A) China storm losses, (B) US tornado losses, (C) AIG loss returns, and (D) Aggregated US market index loss returns, with $\tau_n'=1-1/n$: Estimates $\widetilde{e}^{\star}_{\tau_n'}$ (rainbow curve, asymptotic 95% confidence intervals in dashed blue), $\widehat{e}^{\star}_{\tau_n'}$ (red curve, asymptotic 95% confidence intervals in dashed red) and $\widehat{q}^{\star}_{\tau_n'}$ (gray curve), against the sample fraction $1-\tau_n$. The dashed magenta line is the sample maximum.

Expectile-based ES An expectile-based form of ES as $\mathrm{XTCE}_\tau:=\mathbb{E}(X|X>e_\tau)$ was first considered by [68]. Despite its straightforward interpretability, this tail conditional mean does not fulfill the coherence property in general, but the alternative form $\mathrm{XES}_\tau = \frac{1}{1-\tau}\int_\tau^1 e_t\,\mathrm{d}t$ is coherent (Proposition 2 of [22]). Moreover, for heavy-tailed distributions with tail index $\xi \in (0,1)$ and $\mathbb{E}|X_-| < \infty$, XES_τ is asymptotically equivalent to the intuitive XTCE_τ and proportional to the expectile e_τ since (Proposition 3 of [22]),

$$\mathrm{XES}_\tau \sim \mathrm{XTCE}_\tau \sim \frac{e_\tau}{1-\xi} \quad \text{as} \quad \tau \uparrow 1.$$

Given intermediate and extreme levels $\tau_n, \tau_n' \uparrow 1$ such that $n(1-\tau_n) \to \infty$ and $(1-\tau_n')/(1-\tau_n) \to 0$, this asymptotic connection suggests the extrapolated estimator

$$\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha) = \frac{\overline{e}^{\star}_{\tau_n'}(\alpha)}{1-\widehat{\xi}_{\tau_n}}$$

for $\mathrm{XES}_{\tau_n'}$, obtained by replacing the tail index ξ with its estimator $\widehat{\xi}_{\tau_n}$ and the extreme expectile $e_{\tau_n'}$ with its weighted Weissman-type estimator $\overline{e}^{\star}_{\tau_n'}(\alpha)$. A direct, LAWS-based extrapolated estimator

$$\widehat{\mathrm{XES}}^{\star}_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\xi}_{\tau_n}} \widehat{\mathrm{XES}}_{\tau_n}$$

for $\mathrm{XES}_{\tau_n'}$ follows from the Weissman approximation

$$\frac{\mathrm{XES}_{\tau_n'}}{\mathrm{XES}_{\tau_n}} \sim \frac{e_{\tau_n'}}{e_{\tau_n}} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\xi}$$

by substituting in $\widehat{\xi}_{\tau_n}$ in place of ξ and replacing the intermediate value XES_{τ_n} with its sample counterpart $\widehat{\mathrm{XES}}_{\tau_n} = \frac{1}{1-\tau_n}\int_{\tau_n}^1 \widehat{e}_t\,\mathrm{d}t$. The asymptotic normality of the latter is given in Theorem 6 of [22], and for $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}$ and $\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$ when $\xi < 1/2$, we refer to Theorems 7 and 8 of [22]. In the general case $0 < \xi < 1$, the asymptotic normality of the purely indirect, quantile-based estimator $\overline{\mathrm{XES}}^{\star}_{\tau_n'}(0) = (1-\widehat{\xi}_{\tau_n})^{-1}\,\widetilde{e}^{\star}_{\tau_n'}$ follows as an immediate corollary of Theorem 1.4.2(ii) and Proposition 4 in [22].

Theorem 1.4.3. *(i) Under the conditions of Theorem 1.4.2(i),*

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\widehat{\mathrm{XES}}^{\star}_{\tau_n'}}{\mathrm{XES}_{\tau_n'}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \xi^2\right),$$

and for any $\alpha \in \mathbb{R}$,

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)}{\mathrm{XES}_{\tau_n'}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\frac{\lambda}{1-\rho}, \xi^2\right).$$

(ii) This remains valid for $\alpha = 0$ under the weaker assumptions of Theorem 1.4.2(ii).

Returning to our four data examples, Figure 1.6 displays the plots of the purely indirect, quantile-based estimator $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'} := \overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha = 0)$, its direct, LAWS-based counterpart $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}$ and the Weissman quantile estimator $\widehat{q}^{\star}_{\tau_n'}$, as functions of the sample fraction $1-\tau_n$, for the extreme level $\tau_n' = 1 - 1/n$, along with the 95% asymptotic confidence intervals derived from the asymptotic normality of $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}$ in (A)-(D) and of $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}$ in (D), under the bias condition $\lambda = 0$. The final pointwise estimates $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}$ and $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}$, chosen by using the same eyeballed effective sample fraction threshold for $\widehat{q}^{\star}_{\tau_n'}$, are reported in Table 1.1, with their associated 95% asymptotic confidence intervals.

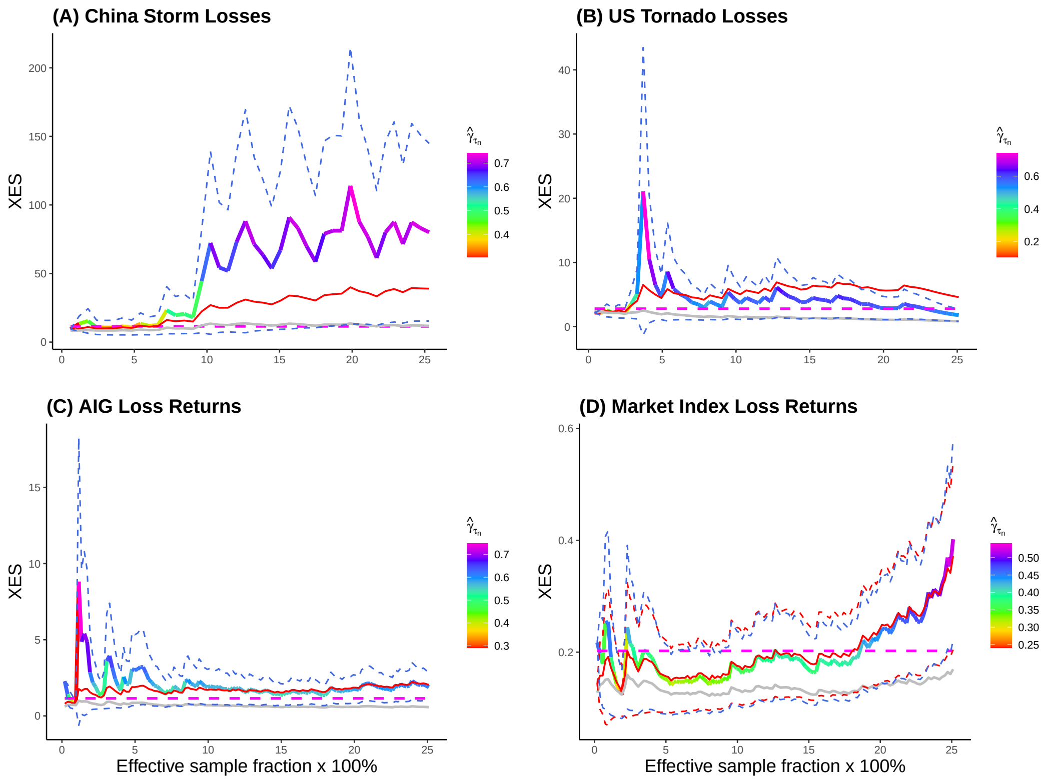

FIGURE 1.6

Extrapolated $\mathrm{XES}_{\tau_n'}$ estimates for (A) China storm losses, (B) US tornado losses, (C) AIG loss returns, and (D) Aggregated US market index loss returns, with $\tau_n' = 1-1/n$: Estimates $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}$ (rainbow curve, asymptotic 95% confidence intervals in dashed blue), $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}$ (red curve, asymptotic 95% confidence intervals in dashed red) and $\widehat{q}^{\star}_{\tau_n'}$ (gray curve), against the sample fraction $1-\tau_n$. The dashed magenta line is the sample maximum.

1.5 Towards multivariate risk assessment: systemic risk measures

The risk measures considered so far only quantify the risk carried by a single random variable. In global economics-oriented examples where there are several players involved, there is valuable information about the dependence between the risk variables of interest that even the joint estimation of several univariate risk measures cannot recover. We illustrate this on the following toy example: consider two random variables X and Y, having respectively a Fréchet distribution with tail index $1/4$ (namely, $\mathbb{P}(X \leqslant x) = \exp(-x^{-4})$ for $x > 0$) and a Pareto distribution with tail index $1/4$ (namely, $\mathbb{P}(Y \leqslant y) = 1 - y^{-4}$ for $y > 1$), and whose dependence structure is given by the Gumbel-Hougaard (or logistic) copula

$$C_\theta(u,v) = \exp\left\{-\left[(-\log u)^\theta + (-\log v)^\theta\right]^{1/\theta}\right\},\ \ u,v \in (0,1),\ \ \theta \geqslant 1.$$

In other words, $\mathbb{P}(X \leqslant x, Y \leqslant y) = C_\theta(\mathbb{P}(X \leqslant x), \mathbb{P}(Y \leqslant y))$. Taking $\theta = 1$ produces a random pair (X,Y) having independent components, and $\theta \to \infty$ yields a perfectly dependent random pair, namely, $F_X(X) = F_Y(Y)$. In this example, the (extreme) univariate risk

Risk level	China storms $n = 166$	US tornadoes $n = 243$	AIG $n = 522$	Market index $n = 522$
$X_{n:n}$	11.44	2.80	1.15	0.20
$\widehat{q}^{\star}_{\tau'_n}$	12.38 (1.77, 22.99)	1.47 (0.42, 2.51)	0.59 (0.28, 0.91)	0.12 (0.07, 0.16)
$\widehat{\mathrm{ES}}^{\star}_{\tau'_n}$	28.09	5.01	1.62	0.19 (0.12, 0.26)
$\widetilde{\mathrm{ES}}^{\star}_{\tau'_n}$	35.10 (5.02, 65.17)	3.44 (0.99, 5.88)	1.35 (0.63, 2.06)	0.18 (0.11, 0.24)
$\widehat{x}^{\star}_{\tau'_n}$	18.85	3.28	1.04	0.14 (0.09, 0.19)
$\widetilde{x}^{\star}_{\tau'_n}$	24.66 (3.53, 45.80)	2.47 (0.71, 4.22)	0.97 (0.46, 1.49)	0.14 (0.09, 0.20)
$\widehat{e}^{\star}_{\tau'_n}$	15.85	2.92	0.71	0.09 (0.05, 0.12)
$\widetilde{e}^{\star}_{\tau'_n}$	18.34 (2.62, 34.05)	1.73 (0.50, 2.97)	0.68 (0.32, 1.04)	0.09 (0.06, 0.13)
$\widehat{\mathrm{XES}}^{\star}_{\tau'_n}$	25.04	5.44	1.64	0.15 (0.09, 0.20)
$\widetilde{\mathrm{XES}}^{\star}_{\tau'_n}$	51.97 (7.44, 96.49)	4.06 (1.17, 6.94)	1.54 (0.72, 2.35)	0.14 (0.09, 0.19)

TABLE 1.1
Estimated extreme risk levels of the China storms, US tornadoes, AIG and aggregated market index datasets at $\tau'_n = 1 - 1/n$ along with 95% asymptotic confidence intervals (available for the direct estimates only in the aggregated market index data analysis where the estimated tail index is found to be $< 1/2$).

measures of X and Y seen so far do not depend on the degree of dependence θ between X and Y, but the MES of X at level $1 - \tau$, defined as

$$\mathrm{MES}_\tau = \mathbb{E}(X | Y > q_\tau(Y)),$$

does depend on θ, as Figure 1.7 shows. Moreover, a stronger degree of positive association between X and Y indeed results in a larger MES risk measure, all other things being equal.

To incorporate the concept of extremal dependence between two random variables into an assessment of risk, suppose that X and Y have continuous survival functions $\overline{F}_X = 1 - F_X$ and $\overline{F}_Y = 1 - F_Y$. We introduce the bivariate survival copula $\overline{C}(u, v) = \mathbb{P}(\overline{F}_X(X) \leqslant u, \ \overline{F}_Y(Y) \leqslant v)$ $(u, v \in [0, 1])$ of X and Y, and we assume the tail dependence condition

$\mathcal{JC}(R)$ There is a function R on $[0, \infty]^2 \backslash \{(\infty, \infty)\}$, with $R(1, 1) > 0$, such that

$$\forall (x, y) \in [0, \infty]^2 \backslash \{(\infty, \infty)\}, \ \lim_{s \to \infty} s\, \overline{C}(x/s, y/s) = R(x, y).$$

This condition imposes the existence of a limiting dependence structure in the joint right tail of X and Y, given by a *tail copula* R (see [65]). It is arguably a minimal assumption when it comes to assessing the dependence structure between extreme value estimators.

In this context, when X is positive and heavy-tailed with tail index $\xi_X \in (0, 1)$,

$$\frac{\mathrm{MES}_\tau}{q_\tau(X)} \to \int_0^\infty R(x^{-1/\xi_X}, 1)\, \mathrm{d}x \quad \text{as} \quad \tau \uparrow 1,$$

see Proposition 1 in [14]. On the basis of this convergence and the Weissman approximation, we get for intermediate and extreme levels $\tau_n, \tau'_n \uparrow 1$ that

$$\mathrm{MES}_{\tau'_n} \sim \frac{q_{\tau'_n}(X)}{q_{\tau_n}(X)} \mathrm{MES}_{\tau_n} \approx \left(\frac{1 - \tau'_n}{1 - \tau_n} \right)^{-\xi_X} \mathrm{MES}_{\tau_n} \quad \text{as} \quad n \to \infty.$$

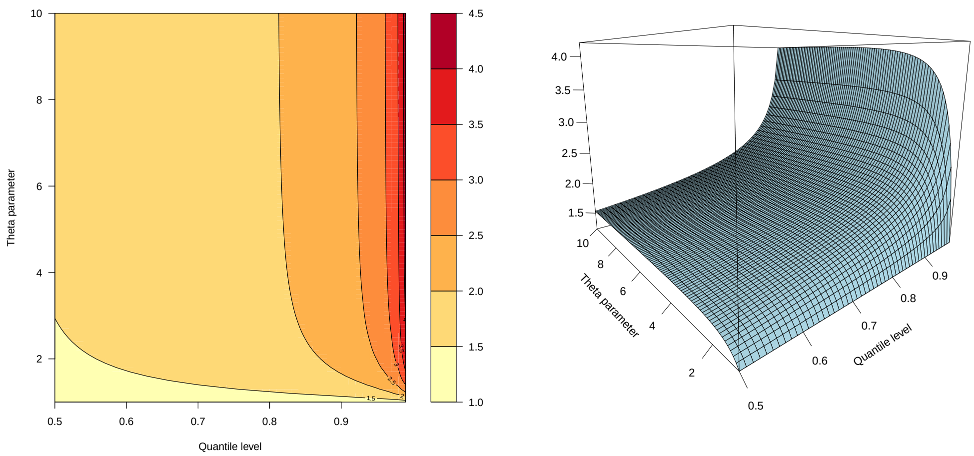

FIGURE 1.7
MES values for a random pair (X, Y) whose first (resp. second) marginal distribution is Fréchet(1/4) (resp. Pareto(1/4)) and whose copula function is the Gumbel-Hougaard copula C_θ, for $\theta \geqslant 1$. The left (resp. right) panel is a heatmap (resp. surface plot) of the values of $\text{MES}_\tau = \mathbb{E}(X|Y > q_\tau(Y))$ as a function of the $Y-$quantile level $\tau \geqslant 1/2$ and the dependence parameter $\theta \geqslant 1$.

Plugging in a $\sqrt{n(1-\tau_n)}$-consistent estimator $\widehat{\xi}_X$ of ξ_X and replacing MES_{τ_n} with its empirical version

$$\frac{\sum_{i=1}^{n} X_i \mathbb{1}\{Y_i > \widehat{q}_{\tau_n}(Y)\}}{\sum_{i=1}^{n} \mathbb{1}\{Y_i > \widehat{q}_{\tau_n}(Y)\}}$$

results in an extreme value estimator for $\text{MES}_{\tau_n'}$. In real-life situations where the profit-loss variable X is the equity return of a financial firm and Y represents an aggregated return measure of a financial market, the random variable X is not necessarily positive, but as shown by [14], its MES is mainly determined by top, and hence positive, realizations of X. This led [14] to propose the intermediate estimator

$$\widehat{\text{MES}}_{\tau_n} = \frac{\sum_{i=1}^{n} X_i \mathbb{1}\{X_i > 0, Y_i > \widehat{q}_{\tau_n}(Y)\}}{\sum_{i=1}^{n} \mathbb{1}\{Y_i > \widehat{q}_{\tau_n}(Y)\}},$$

as well as the corresponding extreme value estimator of $\text{MES}_{\tau_n'}$:

$$\widehat{\text{MES}}^{\star}_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\xi}_X} \widehat{\text{MES}}_{\tau_n}.$$

The asymptotic normality of $\widehat{\text{MES}}^{\star}_{\tau_n'}$ is established in Theorems 1 and 2 of [14], with possibly two different intermediate levels τ_n in $\widehat{\xi}_X$ and $\widehat{\text{MES}}_{\tau_n}$. As evidenced in the asymptotic theory and finite-sample simulation study of [14], selecting different thresholds may result in extrapolated estimators with faster speed of convergence and a better finite-sample performance. Nevertheless, and for the sake of simplicity, we will use in the sequel the same τ_n for both of these estimators. An alternative extreme value estimator for $\text{MES}_{\tau_n'}$ has also been suggested by [14] in their Equation (12) with no asymptotic theory.

Another option studied by [20, 27] is given by the expectile-based form

$$\mathrm{XMES}_\tau = \mathbb{E}(X|Y > e_\tau(Y)),$$

which obeys similar asymptotic approximations as the standard quantile-based form MES_τ. Indeed, assuming that condition $\mathcal{JC}(R)$ holds and X, Y are heavy-tailed with respective tail indices $\xi_X, \xi_Y \in (0,1)$, we have as $\tau \uparrow 1$,

$$\frac{\mathrm{XMES}_\tau}{U_X(1/\overline{F}_Y(e_\tau(Y)))} \to \int_0^\infty R(x^{-1/\xi_X}, 1)\,\mathrm{d}x \quad\text{and}\quad \frac{\mathrm{XMES}_\tau}{\mathrm{MES}_\tau} \to \left(\xi_Y^{-1} - 1\right)^{-\xi_X},$$

with U_X being the tail quantile function of X, see Proposition 2 in [14]. Combining the first convergence result with the Weissman approximation leads to

$$\frac{\mathrm{XMES}_{\tau_n'}}{\mathrm{XMES}_{\tau_n}} \sim \frac{U_X(1/\overline{F}_Y(e_{\tau_n'}(Y)))}{U_X(1/\overline{F}_Y(e_{\tau_n}(Y)))} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\xi_X} \quad \text{as} \quad n \to \infty.$$

A direct, LAWS-based estimator of the extreme value $\mathrm{XMES}_{\tau_n'}$ follows then as

$$\widehat{\mathrm{XMES}}^\star_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\xi}_X} \widehat{\mathrm{XMES}}_{\tau_n},$$
$$\text{where} \quad \widehat{\mathrm{XMES}}_{\tau_n} = \frac{\sum_{i=1}^n X_i \mathbb{1}\{X_i > 0, Y_i > \widehat{e}_{\tau_n}(Y)\}}{\sum_{i=1}^n \mathbb{1}\{Y_i > \widehat{e}_{\tau_n}(Y)\}}.$$

The asymptotic proportionality relationship $\mathrm{XMES}_\tau \sim \left(\xi_Y^{-1} - 1\right)^{-\xi_X} \mathrm{MES}_\tau$ as $\tau \uparrow 1$ also yields the indirect quantile-based estimator

$$\widetilde{\mathrm{XMES}}^\star_{\tau_n'} = \left(\widehat{\xi}_Y^{-1} - 1\right)^{-\widehat{\xi}_X} \widehat{\mathrm{MES}}^\star_{\tau_n'},$$

for a suitable estimator $\widehat{\xi}_Y$ of ξ_Y.

The three extrapolated estimators $\widehat{\mathrm{MES}}^\star_{\tau_n'}$, $\widehat{\mathrm{XMES}}^\star_{\tau_n'}$ and $\widetilde{\mathrm{XMES}}^\star_{\tau_n'}$ have the same asymptotic distribution, dictated by the tail index estimator $\widehat{\xi}_X$, in view of Theorem 2 in [14] and Theorems 4 and 5 in [20] respectively. The result stated below is essentially a consequence of Theorem 4.1 in [27], encompassing the results mentioned above. Introduce the bivariate second-order condition:

$\mathcal{JC}_2(R,\beta,\kappa)$ Condition $\mathcal{JC}(R)$ holds and there are $\beta > \xi_X$ and $\kappa < 0$ such that locally uniformly in $y \in (0,\infty)$,

$$\sup_{x>0} \left| \frac{s\,\overline{C}(x/s, y/s) - R(x,y)}{\min(x^\beta, 1)} \right| = \mathrm{O}(s^\kappa) \quad \text{as} \quad s \to \infty.$$

Theorem 1.5.1. *Suppose that $X > 0$ with probability 1, U_X and U_Y satisfy conditions $\mathcal{C}_2(\xi_X, \rho_X, A_X)$ and $\mathcal{C}_2(\xi_Y, \rho_Y, A_Y)$ with $0 < \xi_X < 1/2$ and $\rho_X < 0$, and condition $\mathcal{JC}_2(R,\beta,\kappa)$ holds. Assume moreover that $\tau_n, \tau_n' \uparrow 1$, with $n(1-\tau_n) \to \infty$, $(1-\tau_n')/(1-\tau_n) \to 0$ and $\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n')) \to \infty$, as well as the bias conditions $n(1-\tau_n)^{1-2\kappa} \to 0$ and $\sqrt{n(1-\tau_n)}|A_X((1-\tau_n)^{-1})|^{\xi_X/(1-\rho_X)-\varepsilon} \to 0$ for some $\varepsilon > 0$, and suppose that $\sqrt{n(1-\tau_n)}(\widehat{\xi}_X - \xi_X) \xrightarrow{\mathrm{d}} Z$ as $n \to \infty$.*

(i) Then

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))} \left(\frac{\widehat{\mathrm{MES}}^\star_{\tau_n'}}{\mathrm{MES}_{\tau_n'}} - 1\right) \xrightarrow{\mathrm{d}} Z.$$

(ii) *Assume* $\xi_Y \in (0,1)$, $\mathbb{E}|Y_-| < \infty$, $\sqrt{n(1-\tau_n)}A_Y((1-\tau_n)^{-1}) \to 0$, $\sqrt{n(1-\tau_n)}/q_{\tau_n}(Y) \to 0$, *and* $\sqrt{n(1-\tau_n)}(\widehat{\xi}_Y - \xi_Y) = \mathrm{O}_{\mathbb{P}}(1)$. *Then*

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\widetilde{\mathrm{XMES}}^{\star}_{\tau_n'}}{\mathrm{XMES}_{\tau_n'}} - 1\right) \xrightarrow{\mathrm{d}} Z.$$

(iii) *Assume* $\xi_Y \in (0,1/2)$ *and there is* $\delta > 0$ *such that* $\mathbb{E}|Y_-|^{2+\delta} < \infty$, *and also that* $\sqrt{n(1-\tau_n)}A_Y((1-\tau_n)^{-1}) \to 0$ *and* $\sqrt{n(1-\tau_n)}/q_{\tau_n}(Y) \to 0$. *Then*

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\widehat{\mathrm{XMES}}^{\star}_{\tau_n'}}{\mathrm{XMES}_{\tau_n'}} - 1\right) \xrightarrow{\mathrm{d}} Z.$$

All convergences remain valid in the case when X *is not necessarily positive, provided* $\mathbb{E}|X_-|^{1/\xi_X} < \infty$ *and* $n(1-\tau_n) = \mathrm{o}((1-\tau_n')^{2\kappa(1-\xi_X)})$ *as* $n \to \infty$.

We apply the three extrapolated estimators $\widehat{\mathrm{MES}}^{\star}_{\tau_n'}$, $\widehat{\mathrm{XMES}}^{\star}_{\tau_n'}$ and $\widetilde{\mathrm{XMES}}^{\star}_{\tau_n'}$ to estimate the two forms of extreme MES for AIG, Citigroup, JPMorgan Chase and Berkshire Hathaway, where for these four examples X refers to the loss return of each company and Y is the aggregated market index considered previously in this chapter (the choice of the frequency of data and time horizon follows the same set-up described in Section 1.2 for AIG and the market index). It should be noted that inference is feasible so far only in the case $\xi_X \in (0,1/2)$ for $\mathrm{MES}_{\tau_n'}$, and in the case $(\xi_X, \xi_Y) \in (0,1/2) \times (0,1)$ for $\mathrm{XMES}_{\tau_n'}$; the condition $\xi_X < 1/2$ does not appear to be satisfied for AIG and Citigroup. The plots of the three extrapolated estimates are graphed in Figure 1.8 as functions of $1-\tau_n$ for $\tau_n' = 1 - 1/n$, along with 95% asymptotic confidence intervals derived from Theorem 1.5.1. We eyeball the effective sample fraction threshold for stability of the estimates and take $1-\tau_n = 13\%$ for AIG, $1-\tau_n = 12\%$ for Citigroup, $1-\tau_n = 9\%$ for JPMorgan Chase, and $1-\tau_n = 17\%$ for Berkshire Hathaway, leading to the final pointwise estimates and corresponding 95% asymptotic confidence intervals reported in Table 1.2.

Financial institution	$\widehat{\xi}_X$	$\widehat{\mathrm{MES}}^{\star}_{\tau_n'}$	$\widehat{\mathrm{XMES}}^{\star}_{\tau_n'}$	$\widetilde{\mathrm{XMES}}^{\star}_{\tau_n'}$
AIG	0.56	0.73	0.71	0.57
Citigroup	0.50	0.97	0.79	0.77
JPMorgan Chase	0.32	0.27 (0.18, 0.37)	0.23 (0.15, 0.31)	0.22 (0.14, 0.30)
Berkshire Hathaway	0.39	0.13 (0.08, 0.17)	0.10 (0.06, 0.13)	0.11 (0.07, 0.15)

TABLE 1.2
Estimates for AIG, Citigroup, JPMorgan Chase and Berkshire Hathaway at $\tau_n' = 1 - 1/n$ with $n = 522$. The second column reports the bias-reduced Hill estimate $\widehat{\xi}_X$ for each institution. The third column reports the quantile-based MES estimates $\widehat{\mathrm{MES}}^{\star}_{\tau_n'}$, and the last two columns report the expectile-based MES estimates $\widehat{\mathrm{XMES}}^{\star}_{\tau_n'}$ and $\widetilde{\mathrm{XMES}}^{\star}_{\tau_n'}$. Each estimate is followed by a 95% asymptotic confidence interval (available when $\widehat{\xi}_X < 1/2$).

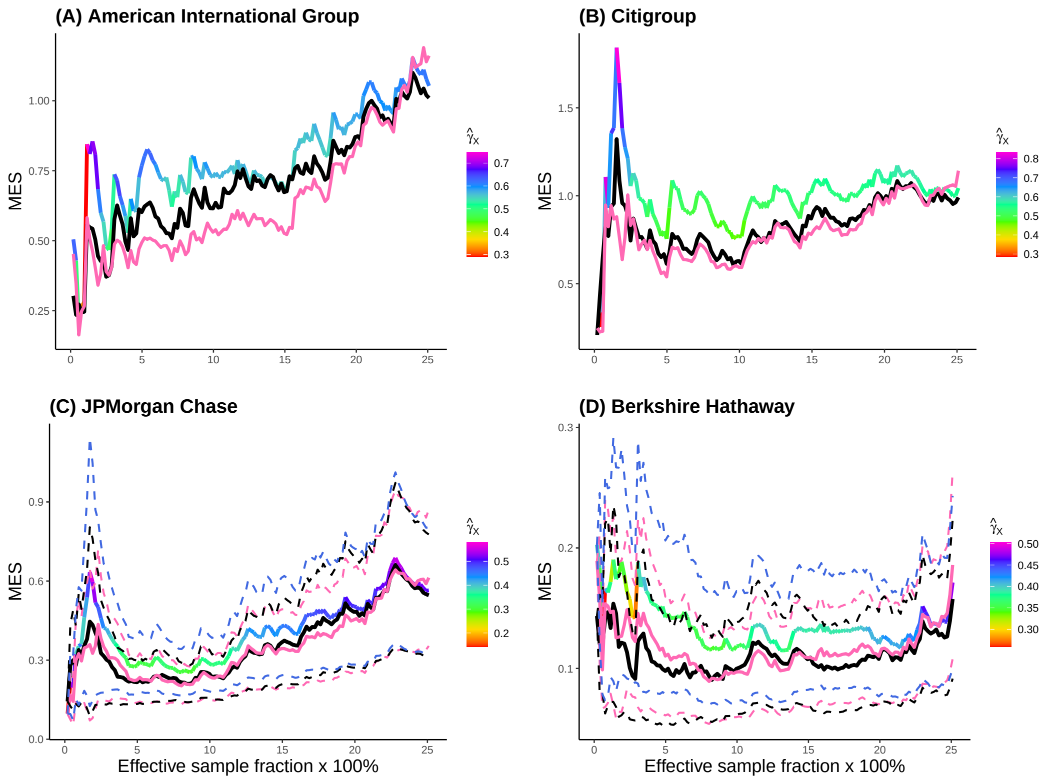

FIGURE 1.8
Extrapolated $\mathrm{MES}_{\tau_n'}$ and $\mathrm{XMES}_{\tau_n'}$ estimates for (A) AIG, (B) Citigroup, (C) JPMorgan Chase, and (D) Berkshire Hathaway, with $\tau_n' = 1 - 1/n$: Estimates $\widehat{\mathrm{MES}}^{\star}_{\tau_n'}$ (rainbow curve, asymptotic 95% confidence intervals in dashed blue), $\widehat{\mathrm{XMES}}^{\star}_{\tau_n'}$ (black curve, asymptotic 95% confidence intervals in dashed black) and $\widetilde{\mathrm{XMES}}^{\star}_{\tau_n'}$ (pink curve, asymptotic 95% confidence intervals in dashed pink), against the sample fraction $1 - \tau_n$.

1.6 Discussion

Much remains to be done if the alternatives to quantiles discussed in this chapter are to be widely used in day-to-day risk management practice. The main question currently open, in our view, is how to carry out accurate inference about extreme risk measures in realistic settings. As far as extreme expectile estimation is concerned, for instance, it appears from the finite-sample results of [20, 22] that the estimators $\widehat{e}^{\star}_{\tau_n'}$ and $\widetilde{e}^{\star}_{\tau_n'}$ suffer from substantial finite-sample bias. It was also noted later by [62] that asymptotic Gaussian inference of extreme expectiles $e_{\tau_n'}$ using these estimators was a difficult question, due to the fact that the asymptotic variances of their Gaussian limiting distributions tend to poorly represent the actual uncertainty in finite samples. One should expect these issues about bias and inference to be more generally present and detrimental when (M)ES or extremiles are used.

Recently, [41] and then [25] have come up with very accurate bias-corrected versions and precise approximations to the empirical variances of $\widehat{e}^{\star}_{\tau_n'}$ and $\widetilde{e}^{\star}_{\tau_n'}$, thus enabling to construct refined bias-reduced and variance-corrected Gaussian confidence intervals for the tail risk measure $e_{\tau_n'}$ in heavy-tailed models. They have provided successive corrections for three sources of approximation errors in the development of the asymptotic normality of both extrapolated estimators, namely (i) the use of the asymptotic connection between extreme expectiles and quantiles while ignoring higher-order error terms, (ii) wrongly neglecting correlations between estimators when the asymptotic behavior of one of them dominates, and (iii) incurring variance distortions by applying the delta-method for linearization purposes. The resulting corrected Gaussian confidence intervals enjoy coverages close to the nominal confidence level even in moderately large samples [25]. The bias-reduced and variance-corrected versions of $\widehat{e}^{\star}_{\tau_n'}$ (see Section 3 of [25]) and $\widetilde{e}^{\star}_{\tau_n'}$ (see Section 4 of [25]), obtained for the aggregated market index and JPMorgan Chase data, at the extreme level $\tau_n' = 1 - 1/n$, are graphed in Figure 1.9, along with corrected associated 95% confidence intervals which better account for statistical uncertainty compared to the original, naive Gaussian confidence intervals we introduced in this chapter. It is noteworthy that at the moment, the bias-reduced version of $\widetilde{e}^{\star}_{\tau_n'}$ appears to be more stable and hence reliable than the one built on $\widehat{e}^{\star}_{\tau_n'}$; subject to further improvements in this direction, the approach developed by [25, 41] could serve as a blueprint for the construction of bias-reduced and variance-corrected asymptotic Gaussian confidence intervals for extreme ES and extremiles.

A broader question of substantial practical interest is to determine when the so-called "direct" estimators considered in this chapter, based on extrapolating intermediate risk measures calculated using the empirical distribution of the data, perform better than "indirect" estimators constructed upon asymptotic proportionality relationships warranted by the extreme value model assumption. A first step towards that goal is a comparison of the asymptotic variances of risk measure estimators at intermediate levels, see Figure 1.10. This was already discussed below Theorem 1.4.1 for expectiles, and it can moreover be seen from Figure 1.10 that, on the range $\xi \in (0, 1/2)$, the empirical ES estimator performs (from the asymptotic variance viewpoint) always worse than its quantile-based counterpart, and that the empirical extremile estimator essentially performs comparably or worse than its quantile-based counterpart. In addition, these four estimators have comparable or higher asymptotic variance than the empirical quantile at the same intermediate level, which suggests that the ES and extremile are in a certain sense more difficult to estimate and infer than tail quantiles. As far as extreme risk measure estimation is concerned, existing literature reveals that:

- For extreme expectile estimation, [20] and [22] provide evidence that the indirect estimator $\widetilde{e}^{\star}_{\tau_n'}$ is superior in the case of non-negative loss distributions, while the direct estimator $\widehat{\widehat{e}}^{\star}_{\tau_n'}$ seems to be best in the case of real-valued profit-loss distributions.
- For tail extremile estimation, [18] provide evidence that the estimator $\widehat{x}^{\star}_{\tau_n'}$ of $x_{\tau_n'}$ is superior in terms of mean-squared error and bias when $\xi < 1/2$ compared to $\widetilde{x}^{\star}_{\tau_n'}$.

Further, large-scale Monte-Carlo studies are required to reach a definitive conclusion and, perhaps, find whether "direct" methods are superior in certain situations regardless to which tail risk measure is estimated. This would be valuable in applications such as financial data analysis where prior information is often available about what kind of model is reasonable. It should also be noted that, as pointed out in Section 1.4, the use of an adaptive expectile estimator, based on a linear combination of the intermediate "direct" estimator and its competing "indirect" analog, is fruitful when the weighting parameter $\alpha \in \mathbb{R}$ is chosen in an optimal way. To the best of our knowledge, this has only been done for tail expectiles in [22, 23]. For tail ES and extremile estimation, the optimal construction of such a linear combination is an open problem.

Last but not least, this introductory chapter to methodologies for extreme risk assessment alternative to quantiles does not deal with more complex real-life situations where

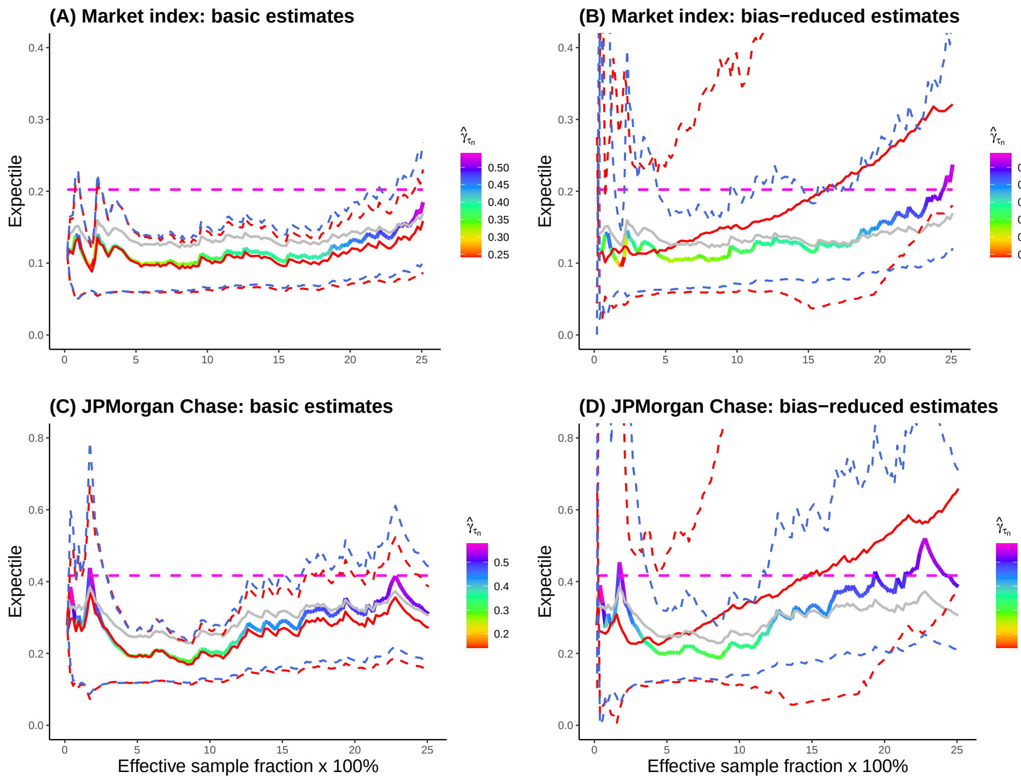

FIGURE 1.9
Extrapolated $e_{\tau_n'}$ estimates for the aggregated US market index (top panels) and JPMorgan Chase (bottom panels) with $\tau_n' = 1 - 1/n$. Left: Basic estimates $\widetilde{e}^{\star}_{\tau_n'}$ (rainbow curve, asymptotic 95% confidence intervals in dashed blue) and $\widehat{e}^{\star}_{\tau_n'}$ (red curve, asymptotic 95% confidence intervals in dashed red), against the sample fraction $1 - \tau_n$. Right: Bias-reduced estimates and variance-corrected confidence intervals. The benchmark Weissman quantile estimates $\widehat{q}^{\star}_{\tau_n'}$ in gray curve and the sample maximum in dashed magenta line.

the data points are serially dependent and/or heterogeneous with or without available covariate information that may be high-dimensional. A fully operational inferential theory that would deal with all these difficulties is not currently available. Various steps have been made towards solving one of the above challenges: the setting of serially dependent and stationary observations, without covariates, has been considered by [27] in the extreme expectile estimation context. The work of [40] gives flexible theory for extreme conditional expectile estimation, including in popular time series, but their approach is highly sensitive to model misspecification, makes the strong assumption of a constant tail index, and their bootstrap scheme is difficult to calibrate. The recent work of [24] provides fully operational extreme conditional quantile and expectile estimators based on $\alpha-$mixing data, although their approach is, by design, limited to low-dimensional covariates; the work of [72] can handle moderately high-dimensional covariates but puts a very strong linear constraint on tail conditional expectiles. Extremile regression for independent, identically distributed and low-dimensional data is considered in [19]. The integration of low-dimensional covariates into extreme MES estimation is discussed in [43, 44]. Dynamic estimation, namely, conditional estimation in time series models where the covariate primarily consists of lags of the target variable with a view on predicting future risk levels, is addressed by [49, 50] in strictly stationary, parametric location-scale models.

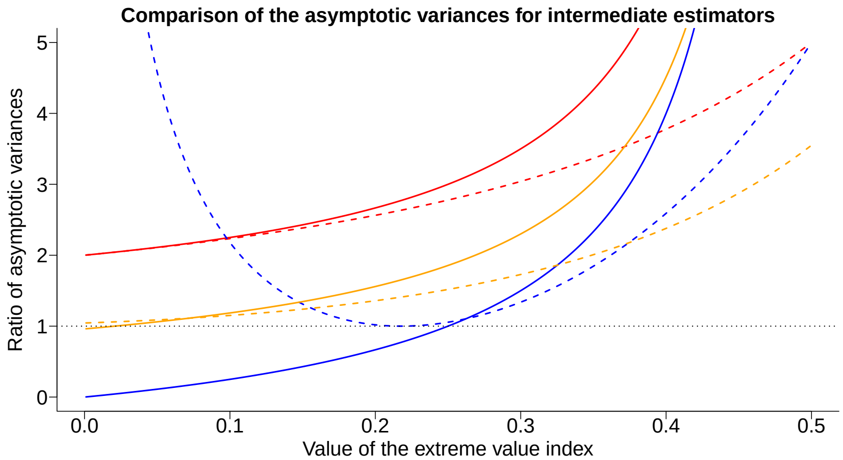

FIGURE 1.10
Comparison of the normalized asymptotic variances $v(\xi)/\xi^2$, where $v(\xi)$ is the asymptotic variance of the empirical ES estimator $\widehat{\text{ES}}_{\tau_n}$ (solid red line), the indirect quantile-based ES estimator $\widetilde{\text{ES}}_{\tau_n}$ (dashed red line), the empirical extremile estimator $\widehat{x}_{\tau_n}$ (solid orange line), the indirect quantile-based extremile estimator $\widetilde{x}_{\tau_n}$ (dashed orange line), the empirical expectile estimator $\widehat{e}_{\tau_n}$ (solid blue line) and the indirect quantile-based expectile estimator $\widetilde{e}_{\tau_n}$ (dashed blue line). The horizontal dotted line is the line $y(\xi) = 1$, corresponding to an asymptotic variance equal to that of the empirical intermediate quantile estimator.

Acknowledgments

This research was supported by the French National Research Agency under the grants ANR-23-CE40-0009 (EXSTA project), ANR-19-CE40-0013 (ExtremReg project), ANR-17-EURE-0010 (EUR CHESS) and ANR-11-LABX-0020-01 (Centre Henri Lebesgue), the TSE-HEC ACPR Chair “Regulation and systemic risks”, and the Chair Stress Test, RISK Management and Financial Steering of the Foundation Ecole Polytechnique.

Bibliography

[1] C. Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518, 2002.

[2] C. Acerbi and B. Székely. Backtesting expected shortfall. *MSCI*, pages 1–34, 2014.

[3] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.

[4] V. V. Acharya, L. H. Pedersen, T. Philippon, and M. Richardson. Measuring Systemic Risk. *The Review of Financial Studies*, 30(1):2–47, 2016.

[5] T. Adrian and M. K. Brunnermeier. CoVaR. *American Economic Review*, 106(7):1705–1741, 2016.

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

[7] F. Bellini and V. Bignozzi. On elicitable risk measures. *Quantitative Finance*, 15:725–733, 2015.

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

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

[10] V. Brazauskas, B. L. Jones, M. L. Puri, and R. Zitikis. Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference*, 138(11):3590–3604, 2008.

[11] J. Breckling and R. Chambers. M-quantiles. *Biometrika*, 75(4):761–772, 1988.

[12] C. Brownlees and R. F. Engle. SRISK: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies*, 30(1):48–79, 2017.

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

[14] J.-J. Cai, J. H. J. Einmahl, L. de Haan, and C. Zhou. Estimation of the marginal expected shortfall: the mean when a related variable is extreme. *Journal of the Royal Statistical Society: Series B*, 77(2):417–442, 2015.

[15] A. Cherny and D. Madan. New measures for performance evaluation. *The Review of Financial Studies*, 22(7):2571–2606, 2009.

[16] R. Cont, R. Deguest, and G. Scandolo. Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, 10(6):593–606, 2010.

[17] J. Daníelsson, P. Embrechts, C. Goodhart, C. Keating, F. Muennich, O. Renault, and H.S. Shin. An academic response to Basel II. *Special paper no. 130, Financial Markets Group, London School of Economics*, 2001.

[18] A. Daouia, I. Gijbels, and G. Stupfler. Extremiles: A new perspective on asymmetric least squares. *Journal of the American Statistical Association*, 114(527):1366–1381, 2019.

[19] A. Daouia, I. Gijbels, and G. Stupfler. Extremile regression. *Journal of the American Statistical Association*, 117(539):1579–1586, 2022.

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

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

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

[23] A. Daouia, S. Girard, and G. Stupfler. ExpectHill estimation, extreme risk and heavy tails. *Journal of Econometrics*, 221(1):97–117, 2021.

[24] A. Daouia, G. Stupfler, and A. Usseglio-Carleve. Inference for extremal regression with dependent heavy-tailed data. *Annals of Statistics*, 51(5):2040–2066, 2023.

[25] A. Daouia, G. Stupfler, and A. Usseglio-Carleve. Bias-reduced and variance-corrected asymptotic Gaussian inference about extreme expectiles. *Statistics and Computing*, 34(4):130, 2024.

[26] A. Daouia, G. Stupfler, and A. Usseglio-Carleve. An expectile computation cookbook. *Statistics and Computing*, 34(3):103, 2024.

[27] A. C. Davison, S. A. Padoan, and G. Stupfler. Tail risk inference via expectiles in heavy-tailed time series. *Journal of Business and Economic Statistics*, 41(3):876–889, 2023.

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

[29] E. H. Deme, S. Girard, and A. Guillou. Reduced-bias estimator of the Conditional Tail Expectation of heavy-tailed distributions. In M. Hallin, D. Mason, D. Pfeifer, and J. Steinebach, editors, *Mathematical Statistics and Limit Theorems*, pages 105–123. Springer, Cham, 2015.

[30] J. Diebolt, A. Guillou, and I. Rached. Approximation of the distribution of excesses through a generalized probability-weighted moment method. *Journal of Statistical Planning and Inference*, 137(3):841–857, 2007.

[31] J. Diebolt, A. Guillou, and P. Ribereau. Asymptotic normality of extreme quantile estimators based on the Peaks-Over-Threshold approach. *Communications in Statistics – Theory and Methods*, 36(5):869–886, 2007.

[32] W. Ehm, T. Gneiting, A. Jordan, and F. Krüger. Of quantiles and expectiles: Consistent scoring functions, Choquet representations, and forecast rankings. *Journal of the Royal Statistical Society: Series B*, 78(3):505–562, 2016.

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

[34] J. El Methni and G. Stupfler. Improved estimators of extreme Wang distortion risk measures for very heavy-tailed distributions. *Econometrics and Statistics*, 6:129–148, 2018.

[35] P. Embrechts and M. Hofert. Statistics and quantitative risk management for banking and insurance. *Annual Review of Statistics and Its Application*, 1:493–514, 2014.

[36] S. Emmer, M. Kratz, and D. Tasche. What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, 18(2):31–60, 2015.

[37] R. F. Engle, E. Jondeau, and M. Rockinger. Systemic Risk in Europe. *Review of Finance*, 19(1):145–190, 2014.

[38] T. Fissler and J. F. Ziegel. Higher order elicitability and Osband's principle. *Annals of Statistics*, 44(4):1680–1707, 2016.

[39] H. Föllmer and A. Schied. *Stochastic Finance – An Introduction in Discrete Time (second edition).* de Gruyter, Berlin, 2004.

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

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

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

[43] Y. Goegebeur, A. Guillou, N. K. L. Ho, and J. Qin. Conditional marginal expected shortfall. *Extremes*, 24(4):797–847, 2021.

[44] Y. Goegebeur, A. Guillou, N. K. L. Ho, and J. Qin. A Weissman-type estimator of the conditional marginal expected shortfall. *Econometrics and Statistics*, 27:173–196, 2023.

[45] J. A. Greenwood, J. M. Landwehr, N. C. Matalas, and J. R. Wallis. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, 15(5):1049–1054, 1979.

[46] A. Guillou, P. Naveau, J. Diebolt, and P. Ribereau. Return level bounds for discrete and continuous random variables. *TEST*, 18(3):584–604, 2009.

[47] H. Gzyl and S. Mayoral. On a relationship between distorted and spectral risk measures. *Revista Economía Financiera*, 15:8–21, 2008.

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

[49] Y. Hoga. Confidence intervals for conditional tail risk measures in ARMA–GARCH models. *Journal of Business & Economic Statistics*, 37(4):613–624, 2019.

[50] Y. Hoga. Limit theory for forecasts of extreme distortion risk measures and expectiles. *Journal of Financial Econometrics*, 20(1):18–44, 2022.

[51] M. C. Jones. Expectiles and M-quantiles are quantiles. *Statistics & Probability Letters*, 20(2):149–153, 1994.

[52] R. Koenker. *Quantile Regression*. Cambridge University Press, New York, 2005.

[53] S. G. Kou, X. Peng, and C. C. Heyde. External risk measures and Basel accords. *Mathematics of Operations Research*, 38(3):393–417, 2013.

[54] V. Krätschmer, A. Schied, and H. Zähle. Qualitative and infinitesimal robustness of tail-dependent statistical functionals. *Journal of Multivariate Analysis*, 103(1):35–47, 2012.

[55] V. Krätschmer, A. Schied, and H. Zähle. Comparative and qualitative robustness for law-invariant risk measures. *Finance and Stochastics*, 18(2):271–295, 2014.

[56] V. Krätschmer, A. Schied, and H. Zähle. Quasi-Hadamard differentiability of general risk functionals and its applications. *Statistics & Risk Modeling*, 32(1):25–47, 2015.

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

[58] C.-M. Kuan, J.-H. Yeh, and Y.-C. Hsu. Assessing value at risk with CARE, the Conditional Autoregressive Expectile models. *Journal of Econometrics*, 150(2):261–270, 2009.

[59] M. Laidi, A. Rassoul, and H. Ould Rouis. Improved estimator of the Conditional Tail Expectation in the case of heavy-tailed losses. *Statistics, Optimization & Information Computing*, 8(1):98–109, 2020.

[60] J. M. Landwehr, N. C. Matalas, and J. R. Wallis. Probability weighted moments compared with some traditional techniques in estimating gumbel parameters and quantiles. *Water Resources Research*, 15(5):1055–1064, 1979.

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

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

[63] C. Philipps. Interpreting expectiles. Available at SSRN: `http://dx.doi.org/10.2139/ssrn.3881402`, 2022.

[64] R. T. Rockafellar and S. Uryasev. Conditional Value-at-Risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471, 2002.

[65] R. Schmidt and U. Stadtmüller. Non-parametric estimation of tail dependence. *Scandinavian Journal of Statistics*, 33(2):307–335, 2006.

[66] I. Steinwart, C. Pasin, R. C. Williamson, and S. Zhang. Elicitation and identification of properties. *JMLR: Workshop and Conference Proceedings*, 35:1–45, 2014.

[67] G. Stupfler. On a relationship between randomly and non-randomly thresholded empirical average excesses for heavy tails. *Extremes*, 22(4):749–769, 2019.

[68] J. Taylor. Estimating Value at Risk and Expected Shortfall using expectiles. *Journal of Financial Econometrics*, 6(2):231–252, 2008.

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

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

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

[72] W. Xu, Y. Hou, and D. Li. Prediction of extremal expectile based on regression models with heteroscedastic extremes. *Journal of Business & Economic Statistics*, 40(2):522–536, 2022.

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

Index