Tail risk in the tail: Estimating high quantiles when a related variable is extreme

Abstract

In this paper we address the problem of high quantile estimation conditional on a related variable being extreme. The problem set-up is of interest in a number applications to evaluate tail risk of a focal variable in the tail of a conditioning variable. A primary example we consider is the assessment of systemic risk in financial markets using a risk measure known as the conditional value-at-risk (CoVaR). The proposed estimator is based on a novel approach to handle the bivariate tail dependence structure through an adjustment factor that can be used in conjunction with univariate high quantile estimation techniques. We establish the asymptotic behavior of the estimator under relatively weak assumptions, and illustrate its performance via simulation studies and a real data example.

Key words: extreme conditional quantile estimation; multivariate extreme value theory; tail dependence function; heavy tails; CoVaR

1 Introduction

In this paper we consider estimation of high quantiles of a random variable given that a related variable is extreme. Assessing tail risk of one focal variable given a tail event in another variable is relevant in a variety of applications. For example, in hydrology, it is of interest to estimate an extreme return level of rainfall at a particular location given that the areal aggregate rainfall exceeds a high threshold. In finance, vulnerability of individual institutions to extremal market movements can be monitored and potentially mitigated by measuring a high quantile, known in finance as Value-at-Risk, of an institution's losses given that the loss on a market index exceeds a high threshold. This resembles the concept of the exposure conditional value-at-risk (CoVaR) in Adrian and Brunnermeier [2016].

Let $Q_X(p)$ denote the $(1-p)$-quantile of random variable X:

$$Q_X(p) = \inf_x \big\{ \mathbb{P}(X > x) \le p \big\}, \qquad p \in (0,1).$$

For two random variables X and Y, we define the conditional quantile at level $\mathbf{p} = (p_1, p_2)$, denoted $Q_{Y|X}(p_2|p_1)$, as the $(1-p_2)$-quantile of Y given that X exceeds its $(1-p_1)$-quantile:

$$\mathbb{P}\big(Y \ge Q_{Y|X}(p_2|p_1) \mid X \ge Q_X(p_1)\big) = p_2, \qquad p_1, p_2 \in (0,1). \tag{1.1}$$

In risk assessment applications, the interest is in values of p_1 and p_2 that are both small such as 1% and 5%. We will hence refer to the functional in (1.1) as the extreme conditional quantile (ECQ).

The estimation of ECQ in (1.1) is a problem of high quantile estimation conditional on the event that a related variable is extreme, i.e., tail risk estimation in the tail. This problem has an additional layer of challenge compared to existing studies such as Cai et al. [2014]. Cai et al. [2014] consider estimation of the mean given an extreme event in another variable, which is a moderate level characteristic of the conditional distribution. To tackle the problem of tail risk estimation in the tail, one needs to model the tail of Y and dependence between X and Y particularly in the joint tail region. We propose an estimator of ECQ based on a novel approach to handle the bivariate tail dependence structure through an adjustment factor that can be used in conjunction

with univariate high quantile estimation techniques.

We explore the connection between the definition of the ECQ in (1.1) and the tail dependence function, assuming existence of the latter. A genuine contribution in this approach is to define and estimate an adjustment factor which captures the impact of the dependence at extremal levels on the ECQ. With this adjustment factor, we can express the ECQ as a quantile at an adjusted tail probability level of the unconditional distribution.

This approach allows us to break the modelling and estimation procedure into the following three steps: (1) estimation of the tail dependence function; (2) computation of the adjustment factor, and (3) univariate high quantiles estimation for the Y variable. Steps (1) and (3) may be handled in a variety of ways. For step (1), we suggest a semi-parametric approach as a way to balance model uncertainty and estimation efficiency in view of data sparsity especially in the joint tail. That is, a suitable parametric model is to be chosen from a number of available models for the tail dependence function, with model parameters estimated using, for instance, the moment estimator of Einmahl et al. [2012]. However, a non-parametric estimator of the tail dependence function could also be used provided a sufficiently large sample size is available for estimation. The adjustment factor in step (2) can then be computed numerically by solving an equation involving the fitted tail dependence function. Finally, for step (3), we adopt the assumption that the focal variable Y is heavy-tailed and use an extreme value non-parametric high quantile estimator (Weissman [1978]).

The available methods for estimation of the tail dependence function, to the best of our knowledge, are only developed in the setting of i.i.d. samples. If the data exhibits features such as serial dependence and non-stationarity, and the interest is in forecasts of the ECQ functional taking into account serial dependence, the above procedure should be preceded by filtering and then applied to residuals. For financial time series, it is common to use ARMA-GARCH filters, whereas environmental data could be filtered using seasonal decomposition models and ARMA processes. We discuss the theoretical validity of such a two-step approach, based on the recent theoretical development in Hoga [2019] for ARMA-GARCH models, and adopt it in our real data analysis.

The rest of the paper is organized as follows. Section 2 describes the proposed methodology

for ECQ estimation. Section 3 provides results on asymptotic theory for the estimator, including consistency and asymptotic normality, and investigates performance of the estimator in finite sample situations through a simulation study. Section 4 is devoted to an application, in which we use the proposed estimator to produce daily CoVaR forecasts for several financial institutions. Conclusion is given in Section 5. The appendices contain the proofs. The data and R code to reproduce numerical results of the paper are available on GitHub https://github.com/menglinzhou/msCoVaR.

2 Methodology

Consider two random variables: Y, a focal random variable, and X, which will be used to define the extreme conditioning event. The key probabilistic assumptions underlying the proposed methodology for the ECQ estimation include the existence of a non-zero upper tail dependence function R, and that the focal random variable Y has a regular varying right tail: $1 - F_Y \in RV_{-1/\gamma}$ for some $\gamma > 0$. Section 2.1 provides the preliminaries for these two key assumptions. Note that no distributional assumptions are made on random variable X, apart from having a continuous marginal distribution.

Section 2.2 explains the intuition for constructing the estimator, in particular, our main methodological innovation: the introduction of the adjustment factor and its approximation. Section 2.3 details the estimator based on the intuition.

2.1 Preliminaries

We assume regular variation for the right tail of the focal variable Y as follows. A distribution function (df) F on $\mathbb{R}$ with an infinite upper endpoint is <u>regularly varying</u> with index $\alpha > 0$, written as $1 - F \in RV_{-\alpha}$, if for all $x > 0$

$$\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}.$$

Examples of distributions with a regularly varying upper tail include Pareto-like distributions: $1 - F(x) \sim cx^{-\alpha}$, as $x \to \infty$ for some $\alpha, c > 0$. Univariate regular variation characterizes the

maximum domain of attraction of the Fréchet distribution [Gnedenko, 1943]. The reciprocal of the index of regular variation $\gamma = 1/\alpha$ is referred to as the tail index with larger values corresponding to heavier tails.

To model the dependence between Y and X, we focus on the tail dependence properties of the random vector (X, Y). There exist a number of ways to describe the extremal dependence structure of a random vector, including the exponent measure and stable tail dependence function; see, e.g., de Haan and Ferreira [2006]. In our proposed approach, we make use of the (upper) tail dependence function, also known as the tail copula.

Definition 1. *Consider a random vector (X, Y) with joint df F and continuous margins F_X, F_Y. The df F is said to have the (upper) tail dependence function R if for all $x, y > 0$, the following limit exists:*

$$\lim_{u \to 0} \frac{\mathbb{P}\{F_X(X) \geq 1 - ux, F_Y(Y) \geq 1 - uy\}}{u} = R(x, y). \tag{2.1}$$

Note that $R(1, 1)$ is also called the upper tail dependence coefficient [Joe, 1997] and is a popular measure of extremal dependence and risk contagion in finance [McNeil et al., 2005]. The case $R(1, 1) = 0$ is referred to as tail independence, and otherwise we have tail dependence. In the proposition below, we summarize several notable properties of the tail dependence function R; for details, see de Haan and Ferreira [2006], Chapter 6.1.5.

Proposition 2.1. *Let R denote an upper tail dependence function.*

1) R is continuous.

2) R is monotonically non-decreasing in each component.

3) $0 \leq R(x, y) \leq x \wedge y$.

4) R is homogeneous of order 1: $R(tx, ty) = tR(x, y)$ for any $t > 0$.

2.2 Intuition

Let $\mathbf{p} = (p_1, p_2)$ denote probability levels for the ECQ in (1.1). We introduce an adjustment factor $\eta_{\mathbf{p}}$ defined as

$$\eta_{\mathbf{p}} = \frac{\mathbb{P}\big(Y \geq Q_{Y|X}(p_2|p_1)\big)}{\mathbb{P}\big(Y \geq Q_{Y|X}(p_2|p_1) \mid X \geq Q_X(p_1)\big)}, \qquad p_1, p_2 \in (0,1). \tag{2.2}$$

It then follows that $\mathbb{P}\big(Y \geq Q_{Y|X}(p_2|p_1)\big) = p_2\eta_{\mathbf{p}}$ and hence the ECQ is related to a quantile of the unconditional distribution of Y via

$$Q_{Y|X}(p_2|p_1) = Q_Y(p_2\eta_{\mathbf{p}}) \tag{2.3}$$

with the adjusted quantile level $p_2\eta_{\mathbf{p}}$. If X and Y are independent, then the ECQ coincides with the quantile of Y at the same level and $\eta_{\mathbf{p}} = 1$. In the case of positive quadrant dependence, i.e., when $\mathbb{P}(X \geq x, Y \geq y) \geq \mathbb{P}(X \geq x)\mathbb{P}(Y \geq y)$ for $x, y \in \mathbb{R}$, we have $\eta_{\mathbf{p}} \leq 1$ and the ECQ is equal to the quantile at a higher confidence level determined by $\eta_{\mathbf{p}}$.

Going back to the definition of the ECQ in (1.1) and using (2.3), we have

$$\frac{\mathbb{P}\big(X \geq Q_X(p_1),\ Y \geq Q_{Y|X}(p_2|p_1)\big)}{p_1} = p_2.$$

Since the interest lies in small values of risk measure levels p_1 and p_2, the ratio above can be approximated using the tail dependence function in (2.1) as follows:

$$\frac{\mathbb{P}\big(F_X(X) \geq 1 - p_1,\ F_Y(Y) \geq 1 - \eta_{\mathbf{p}}\frac{p_2}{p_1}p_1\big)}{p_1} \approx R\Big(1, \eta_{\mathbf{p}}\frac{p_2}{p_1}\Big)$$

for p_1 and p_2 sufficiently close to 0. This suggests a possibility of obtaining the adjustment factor $\eta_{\mathbf{p}}$ in (2.2) with an approximate version, denoted $\eta_{\mathbf{p}}^*$, which is defined implicitly via

$$R\Big(1, \eta_{\mathbf{p}}^*\frac{p_2}{p_1}\Big) = p_2. \tag{2.4}$$

In situations where tail dependence function provides a good approximation of the dependence structure in the tail region, we expect $\eta_{\mathbf{p}}$ and $\eta_{\mathbf{p}}^*$ to be close. We prove this formally in Lemma A.2 and examine this approximation in simulation studies in Section 3.3.

As the tail dependence function is monotonically non-decreasing in each coordinate (see Proposition 2.1), it follows that $R(1, \eta_{\mathbf{p}}^*\frac{p_2}{p_1})$ is increasing from zero to $R(1,1)$ for values of $\eta_{\mathbf{p}}$ ranging from

6

zero to p_1/p_2. Hence, a unique solution $\eta_{\mathbf{p}}^*$ to equation (2.4) exists provided that $p_1 < 1/R(1,1)$. In applications, p_1 will typically be taken to be small and so generally it will be possible to find the solution in the tail dependence case.

When $\eta_{\mathbf{p}}^*$ and $\eta_{\mathbf{p}}$ are sufficiently close, the ECQ value, equivalent to the unconditional quantile at a more extreme level $p_2\eta_{\mathbf{p}}$ based on (2.3), can be approximated using the regular variation of $1 - F_Y$ as follows:

$$Q_{Y|X}(p_2|p_1) \approx Q_Y(p_2\eta_{\mathbf{p}}^*) \approx (\eta_{\mathbf{p}}^*)^{-\gamma}Q_Y(p_2) \qquad \text{for } p_1, p_2 \text{ close to zero,} \qquad (2.5)$$

where the second approximation uses regular variation of Q_Y (at zero) with index $-\gamma$; see (3.1) below. The assumption of regular variation allows one to extrapolate from a less extreme probability level (p_2 here) to a more extreme one ($p_2\eta_{\mathbf{p}}^*$) with the extrapolation factor that depends on the index of regular variation.

The above expression will be used as a basis for constructing the estimator of the ECQ, by estimating each component on the right hand side.

2.3 Estimation

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ denote i.i.d. copies of random vector (X, Y) from a distribution satisfying the two assumptions that a non-zero upper tail dependence function $R(\cdot, \cdot)$ exists and the random variable Y has a regularly varying right tail. We propose the following estimator of $Q_{Y|X}(p_2|p_1)$:

$$\widehat{Q}_{Y|X}(p_2|p_1) = (\hat{\eta}_{\mathbf{p}}^*)^{-\hat{\gamma}}\widehat{Q}_Y(p_2), \qquad (2.6)$$

with the estimators of each component as follows. The tail index γ of the distribution of Y is estimated using the Hill estimator [Hill, 1975]:

$$\hat{\gamma} = \frac{1}{k_1}\sum_{i=1}^{k_1} \log Y_{n,n-i+1} - \log Y_{n,n-k_1}, \qquad (2.7)$$

where $Y_{n,1} \le Y_{n,2} \le \cdots \le Y_{n,n}$ are the order statistics of the sample $Y_1, \ldots, Y_n$. The choice of the intermediate sequence k_1 can be automatically decided by a two-step subsample bootstrap method in Danielsson et al. [2001].

7

We use the extreme quantile estimator in [Weissman, 1978] for the unconditional quantile:

$$\widehat{Q}_Y(p_2) = Y_{n,n-k_2} \left(\frac{k_2}{np_2} \right)^{\hat{\gamma}}, \tag{2.8}$$

where the choice of the sample fraction k_2 typically aligns with that of k_1.

Finding $\hat{\eta}^*_{\mathbf{p}}$ in (2.6) requires an estimate of the tail dependence function. While a number of non-parametric estimators have been proposed in the literature, a parametric assumption on the form of R will lead to efficiency gains in light of data sparsity in the tail region as well as will facilitate computation of an estimate of $\eta^*_{\mathbf{p}}$. Assuming a parametric model for the tail dependence function $R(\cdot, \cdot) = R(\cdot, \cdot; \boldsymbol{\theta})$, the parameter $\boldsymbol{\theta}$ can be estimated using one of the methods available in the literature for this estimation problem. Coles and Tawn [1991] and Joe et al. [1992] apply maximum likelihood method, while Ledford and Tawn [1996] and Smith [1994] use a censored likelihood approach. Einmahl et al. [2008] point out that these likelihood-based estimation methods require smoothness of the partial derivatives of the tail dependence function. As an alternative, they propose an estimator based on the method-of-moments for dimension two, which requires a smaller set of conditions. We adopt the method-of-moments estimator (M-estimator) proposed in Einmahl et al. [2008]. This M-estimator has been extended in Einmahl et al. [2012] to be used in arbitrary dimensions and its consistency and asymptotic normality hold under weak conditions. Note that any estimator for the parametric tail dependence function with a valid asymptotic theory can serves this purpose. Our choice here is just an example. Another example is the weighted least squares estimator in Einmahl et al. [2018].

We conclude this subsection with the definition of the M-estimator of the tail dependence function. Let R_i^X and R_i^Y denote, respectively, the rank of X_i among $X_1, ..., X_n$ and the rank of Y_i among $Y_1, .., Y_n$ for $i \in \{1, ..., n\}$. A nonparametric estimator of the bivariate upper tail dependence function R is given by:

$$\hat{R}_n(x, y) := \frac{1}{m} \sum_{i=1}^{n} \mathbb{1} \left\{ R_i^X \geq n + \frac{1}{2} - mx, R_i^Y \geq n + \frac{1}{2} - my \right\}, \tag{2.9}$$

where $m = m_n \in \{1, ..., n\}$ is an intermediate sequence.

Assume that the function R belongs to some parametric family $\{R(\cdot, \cdot; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$, where $\Theta \subset \mathbb{R}^q$ $(q \geq 1)$ is the parameter space. Let $g = (g_1, ..., g_q)^T : [0,1]^2 \to \mathbb{R}^q$ be a vector of integrable functions. Define function $\varphi : \Theta \to \mathbb{R}^q$ as:

$$\varphi(\boldsymbol{\theta}) := \int\int_{[0,1]^2} g(x,y) R(x,y;\boldsymbol{\theta}) dx dy. \tag{2.10}$$

Let $\boldsymbol{\theta}_0$ denote the true value of parameter $\boldsymbol{\theta}$. The M-estimator $\hat{\boldsymbol{\theta}}$ is defined as the minimizer of the function (Einmahl et al. [2012])

$$S_{m,n}(\boldsymbol{\theta}) = \left\| \varphi(\boldsymbol{\theta}) - \int\int_{[0,1]^2} g(x,y) \hat{R}_n(x,y) dx dy \right\|^2, \tag{2.11}$$

where $|| \cdot ||$ is the Euclidean norm, and $\hat{R}_n(x,y)$ is the nonparametric estimator of R in (2.9). The choice of m and test function g is discussed further in Section 3.3.

Once we have $\hat{\boldsymbol{\theta}}$, $\hat{\eta}_{\mathbf{p}}^*$ in (2.6) can obtained by solving

$$R\left(1, \hat{\eta}_{\mathbf{p}}^* \frac{p_2}{p_1}; \hat{\boldsymbol{\theta}}\right) = p_2. \tag{2.12}$$

3   Asymptotic theory and simulations

We validate the proposed ECQ estimator by showing its asymptotic properties including consistency and asymptotic normality, as well as its finite sample performance through a simulation study.

3.1   Consistency

In this section we state consistency of the proposed ECQ estimator, and begin by imposing the necessary assumptions to guarantee this result.

Firstly, we present three assumptions that are necessary for consistency of the high quantile estimator in (2.8); see, e.g., Theorem 4.3.8 in de Haan and Ferreira [2006]. The first one is the second order condition on the distribution function of Y, the second one is for the two intermediate sequences $k_j = k_j(n)$, $j = 1, 2$, used in the estimator of quantile in (2.8), and the third one is about the probability levels $p_1 = p_1(n)$ and $p_2 = p_2(n)$.

Denote the quantile function $U_Y = (1/1 - F_Y)^{\leftarrow}$, where $\cdot^{\leftarrow}$ is the left-continuous inverse. Then clearly $Q_Y(p) = U_Y(1/p)$. In addition, the regular variation of $1 - F_Y$ is equivalent to the regular variation of U_Y: for all $x > 0$,

$$\lim_{t \to \infty} \frac{U_Y(tx)}{U_Y(t)} = x^{\gamma}. \tag{3.1}$$

Condition A. Assume that there exist a constant $\rho < 0$ and an eventually positive or negative function $A(t)$ such that as $t \to \infty$, $A(t) \to 0$ and for all $x > 0$,

$$\lim_{t \to \infty} \frac{\frac{U_Y(tx)}{U_Y(t)} - x^{\gamma}}{A(t)} = x^{\gamma} \frac{x^{\rho} - 1}{\rho}. \tag{3.2}$$

This condition quantifies the speed of convergence in the regular variation of U_Y in (3.1).

Condition B. Assume that the intermediate sequences satisfy that as $n \to \infty$

$$k_j \to \infty, \ k_j/n \to 0 \ \text{and} \ \sqrt{k_j} A(n/k_j) \to \lambda_j \in \mathbb{R}, \quad j = 1, 2. \tag{3.3}$$

This condition is often assumed in extreme value statistics to guarantee finite asymptotic bias.

Condition C. Assume that, as $n \to \infty$, $p_1 = p_1(n) \to 0$, $p_2 = p_2(n) \to 0$, $p_1/p_2 \to c$ for some $c \in (0, \infty)$ and the probability level p_1 is compatible with the intermediate sequences k_j, $j = 1, 2$ as follows:

$$\frac{k_2}{np_1} \to \infty \ \text{and} \ \frac{\sqrt{k_1}}{\log(k_2/(np_1))} \to \infty, \quad \text{as } n \to \infty. \tag{3.4}$$

This condition provides both upper and lower bounds for the probability levels $p_1 = p_1(n)$ and $p_2 = p_2(n)$. Both bounds are standard assumptions for high quantile estimation in extreme value statistics.

Next, we present the assumptions from Theorem 4.1 in Einmahl et al. [2012], which guarantee the existence, uniqueness and consistency of M-estimator $\hat{\boldsymbol{\theta}}$.

Condition D. (i) The function φ defined in (2.10) is homeomorphism from $\Theta \to \mathbb{R}^q$ and there exists $\epsilon_0 > 0$ such that the set $\{\boldsymbol{\theta} \in \Theta : ||\boldsymbol{\theta} - \boldsymbol{\theta}_0|| \leq \epsilon_0\}$ is closed; (ii) $\boldsymbol{\theta}_0$ is in the interior of the parameter space Θ, φ is twice continuously differentiable and the total derivative of φ at $\boldsymbol{\theta}_0$ is of full rank.

Last but not least, we impose two conditions on the tail dependence function R in (2.1). The first one aims at controlling the speed of convergence to the limit in (2.1) by a power function and the second one is about the partial derivative of R: $R_2(x, y; \boldsymbol{\theta}) := \partial R(x, y; \boldsymbol{\theta})/\partial y$.

Condition E. There exists a constant $\tilde{\rho} > 1$, such that as $u \to 0$, uniformly for all $(x, y) \in [0, 1]^2 \setminus \{(0, 0)\}$

$$\frac{1}{u} \mathbb{P}\{F_X(X) > 1 - ux, F_Y(Y) > 1 - uy\} - R(x, y) = O(u^{\tilde{\rho}}). \tag{3.5}$$

Note that a similar condition has been assumed for the M-estimator for $\boldsymbol{\theta}$, see assumption (C1) in Einmahl et al. [2012].

Condition F. For all $\boldsymbol{\theta} \in \Theta$, the partial derivative $R_2(x, y; \boldsymbol{\theta})$ is continuous with respect to y in the neighborhood of $(1, 0; \boldsymbol{\theta})$ and $R_2(1, 0; \boldsymbol{\theta}) > 0$.

Notice that we are going to handle $R\left(1, \eta_{\mathbf{p}}^* \frac{p_2}{p_1}\right) = p_2$ as in (2.4). As $p_1 \to 0$ (and $p_2 \to 0$), Condition F ensures that $\eta_{\mathbf{p}} \to 0$ with the same speed as p_1 and p_2 for all tail dependence functions in the parametric family.

The following theorem shows consistency of the ECQ estimator defined in (2.6). The proof is given in Appendix A.

Theorem 3.1. *Assume that Condition A-Condition F hold. Then, as $n \to \infty$,*

$$\frac{\widehat{Q}_{Y|X}(p_1, p_2)}{Q_{Y|X}(p_1, p_2)} \xrightarrow{\mathbb{P}} 1.$$

3.2 Asymptotic normality

In this section we state asymptotic normality of the proposed ECQ estimator under some stronger conditions than those assumed for consistency. We present the conditions first, using the apostrophe sign to indicate the stronger version of a corresponding condition presented in Section 3.1. Conditions A' and D' are identical to conditions A and D, respectively.

Define

$$s_n := \min\left(\sqrt{k_2}, \frac{\sqrt{k_1}}{\log(k_2/(np_1p_2))}\right).$$

We shall show that s_n is the speed of convergence of the ECQ estimator.

Condition B'. Condition B holds. In addition, $k_2/m \to 0$ as $n \to \infty$, where m is the number of order statistics used in the M-estimator $\hat{\boldsymbol{\theta}}$.

This condition ensures that the estimation error for $\boldsymbol{\theta}$ is of a lower order compared to those of the other components and consequently its asymptotic behavior does not affect the asymptotic behavior of the ECQ estimator.

Note that, if a nonparametric estimator of the tail dependence function is used in place of the parametric estimator, stronger restrictions on the k_2 sequence than those required in Condition B' will be needed, due to a slower speed of convergence for the nonparametric estimator.

Condition C'. As $n \to \infty$, $p_1 = p_1(n) \to 0$, $p_2 = p_2(n) \to 0$, $p_1/p_2 \to c$ for some $c \in (0, \infty)$ and the probability levels p_1 and p_2 are compatible with the intermediate sequences k_j, $j = 1, 2$ as follows:

$$\frac{k_2}{np_1} \to \infty \quad \text{and} \quad \frac{\sqrt{k_1}}{\log(k_2/(np_1p_2))} \to \infty, \quad \text{as } n \to \infty. \tag{3.6}$$

This condition implies Condition C and requires a slightly more restrictive lower bound for $p_1 = p_1(n)$ and $p_2 = p_2(n)$. It guarantees that $s_n \to \infty$ as $n \to \infty$.

Condition E'. Condition E holds. In addition, as $n \to \infty$, $m = o(n^{\frac{2\tilde{\rho}}{2\tilde{\rho}+1}})$.

This condition is identical to assumption (C2) in Einmahl et al. [2012], which ensures the asymptotic normality of $\hat{\boldsymbol{\theta}}$ with no asymptotic bias.

Condition F'. Condition F holds. In addition, for each given $\boldsymbol{\theta} \in \Theta$, there exists a function $\breve{\rho}(\boldsymbol{\theta}) > 0$, continuous at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, such that as $y \to 0$,

$$\frac{R(1, y; \boldsymbol{\theta})}{y} - R_2(1, 0; \boldsymbol{\theta}) = O(y^{\breve{\rho}(\boldsymbol{\theta})}). \tag{3.7}$$

In addition, there exists a constant $\varepsilon > 0$ such that $p_1^{\min(\breve{\rho}(\boldsymbol{\theta}_0) - \varepsilon, \tilde{\rho} - 1)} s_n \to 0$ as $n \to \infty$.

This condition is an additional restriction on the upper bound for $p_1 = p_1(n)$ (and $p_2 = p_2(n)$).

Condition G'. The function $R_2(1,0;\boldsymbol{\theta})$ is $1-$Lipschitz continuous in the neighborhood of $\boldsymbol{\theta}_0$.

We remark that all of the above conditions are compatible. For example, one can choose $m = n^\mu$ with $0 < \mu < \frac{2\tilde{\rho}}{2\tilde{\rho}+1}$ and $k_1 = k_2 = k(n) = n^\kappa$ with $0 < \kappa < \mu$. Then the speed of convergence is $s_n = \frac{\sqrt{k}}{\log \frac{k}{np_1p_2}}$ provided that our ECQ estimator is applied to a probability level $p_2 = p_2(n) = n^{-\zeta}$ for $\zeta > \max\left(\frac{\kappa}{2\min(\check{\rho}(\boldsymbol{\theta}_0),\tilde{\rho}-1)}, 1-\kappa\right)$.

Consider the ECQ estimator defined in (2.6). Here, $\hat{\gamma}$ is estimated by the Hill estimator; $\widehat{Q}_Y(p)$ is estimated using (2.8) and $\hat{\eta}_{\mathbf{p}}^*$ is estimated with (2.12). The following theorem shows asymptotic normality of the ECQ estimator in (2.6). The proof is given in Appendix B.

Theorem 3.2. *Assume that Conditions A'-G' hold. In addition, assume that, as $n \to \infty$,*
$$\frac{\sqrt{k_2}}{\sqrt{k_1}/\log(k_2/np_1p_2)} \to r \in [0,+\infty]. \quad \text{Then, as } n \to \infty,$$

$$s_n\left(\frac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)} - 1\right) \xrightarrow{d} \mathcal{N}(\text{bias, variance}),$$

where

$$bias = \frac{\lambda_1}{1-\rho}\min(r,1) \quad and \quad variance = \gamma^2(1+\min(r^2,1/r^2)).$$

3.3 Simulation studies

Several simulation studies are conducted in order to assess finite sample properties of the proposed ECQ estimator.

To evaluate the ECQ estimator in (2.6) for a given value of level $\mathbf{p} = (p_1, p_2)$, we need to obtain estimates of the tail index parameter γ, adjustment factor $\eta_{\mathbf{p}}^*$ through the estimate of the tail dependence function, and $Q_Y(p_2)$, the $(1 - p_2)$-quantile of the distribution of Y. These three components will naturally all have an impact on the performance of the ECQ estimator. Thus, we also report the behavior of the individual components that comprise the ECQ estimator to better understand main sources of its bias and estimation uncertainty.

The assessment is based on 1000 Monte Carlo replications at ECQ level $\mathbf{p} = (0.05, 0.05)$. The i.i.d. samples of size $n = 3000$ are simulated from the following four probability distributions; see Supplementary Material Section S1 for details:

(1) Bivariate logistic distribution with dependence parameter $\theta \in (0, 1]$ and unit Fréchet margins.

(2) Bivariate Hüsler-Reiss (HR) distribution (Hüsler and Reiss [1989]) with dependence parameter $\theta > 0$ and unit Fréchet margins.

(3) Bivariate asymmetric logistic distribution with dependence parameter $\theta \in (0, 1]$ and asymmetry parameters $\phi_1, \phi_2 \in [0, 1]$ and unit Fréchet margins.

(4) Standard bivariate t distribution with $\nu > 0$ degrees of freedom and correlation parameter $\rho \in (-1, 1)$.

The first three models belong to the class of bivariate extreme value distributions, and hence the tail dependence function for these distributions gives the exact representation of the underlying dependence structure. The bivariate t distribution is multivariate regularly varying, and consequentially lies in the domain of attraction of a bivariate extreme value distribution with Fréchet margins (see, e.g., Example 5.21 in Resnick [1987]). In this case, the tail dependence function approximates the dependence structure of the underlying distribution in the joint tail region. We provide expressions for the tail dependence function for each model in the Supplementary Material.

The settings of the simulation studies are summarized in Table 1. In particular, for each model used for data generation, we indicate the values of the model parameters and give specifications for sample fractions k_1 and k_2 and for the M-estimation of the parameters of the tail dependence function including the value of parameter m and test function g. The values of k_1 and k_2 are chosen so as to minimize the mean squared error (MSE) of the estimators of tail index γ and quantile $Q_Y(p_2)$, respectively. The choice of tuning parameter m and function $g(x, y)$ for carrying out M-estimation of the parameters of the tail dependence function is based on the guidance given in Einmahl et al. [2012]. As the M-estimator is sensitive to the value of m, its choice in simulation studies is made on the basis of the behavior of the bias and MSE. Larger values of m lead to increase in the absolute value of the bias. However, in most cases, m can be chosen so as to minimize the MSE. The choice of function $g(x, y)$ does not exert much influence on the M-estimator of the tail

dependence function and can be taken to have a simple form to facilitate computations. Our specific choices are presented in Table 1.

Table 1: The setting of the simulation studies including distributions for data generation and their parameter values, and specifications for ECQ estimation: k_1 (for tail index estimation), k_2 (for high quantile estimation), and m and test function $g(x, y)$ (for estimation of parameters of the tail dependence function).

Model	Parameters	k_1	k_2	m	$g(x, y)$
Logistic	$\theta = 0.6$	360	360	270	$g(x, y) = 1$
HR	$\theta = 2.5$	420	410	420	$g(x, y) = x$
Asymmetric logistic	$(\theta, \psi_1, \psi_2) = (0.6, 0.5, 0.8)$	410	410	240	$g(x, y) = (1, x, 2x + 2y)^T$
Bivariate t	$(\nu, \rho) = (3, 0.6)$	30	150	90	$g(x, y) = (x, x + y)^T$

The summary statistics of the ECQ estimates are reported in Table 2. The first row gives the true values of $Q_{Y|X}(p_2|p_1)$ under the various considered models. As both mean and median of the estimates exceed the associated true ECQ values, these results reveal a tendency of the proposed estimator to overestimate the true value. However, approximate 95% confidence intervals based on asymptotic normality of the sample mean do cover the true values. From the applied perspective, the proposed estimation procedure offers a conservative estimator of risk as measured by the conditional quantile.

We have further investigated the sources of bias and variance of the ECQ estimator. In particular, the panel in Table 2 labelled "True γ" presents summary statistics of ECQ estimates with the tail index kept at its true value, rather than being estimated. Here we observe a substantial bias and variance reduction. The bottom two panels show results based on the true values of the approximate adjustment factor $\eta_{\mathbf{p}}^*$ and the exact adjustment factor $\eta_{\mathbf{p}}$. There is little difference in ECQ estimates when the true adjustment factor $\eta_{\mathbf{p}}$ is used instead of the asymptotic approximation $\eta_{\mathbf{p}}^*$. And while, as expected, the use of true values of either $\eta_{\mathbf{p}}$ or $\eta_{\mathbf{p}}^*$ reduces both the bias and

standard deviation of the ECQ estimator, the reductions are modest compared to those of the tail index parameter. Hence, we can conclude that it is the estimator of γ that is largely responsible for the bias and variability of the proposed ECQ estimator.

Table 2: Summary statistics of ECQ estimates at level $\mathbf{p} = (0.05, 0.05)$ for simulation settings specified in Table 1. The first row gives the true value of the ECQ under each model. The bottom three panels correspond to ECQ estimates when the specified component is held at the true value rather than being estimated.

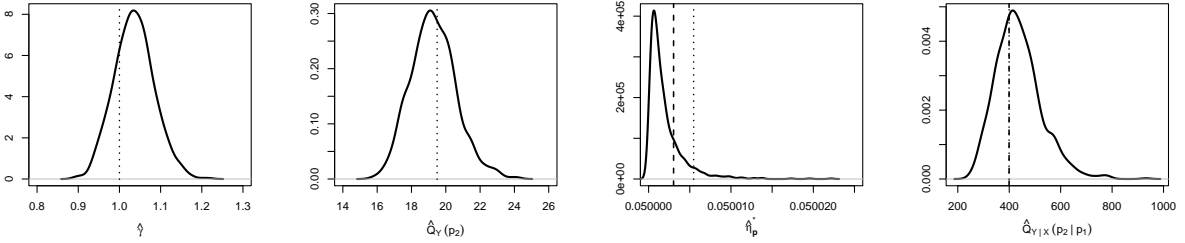|	logistic	HR	asymmetric logistic	bivariate t
$Q_{Y|X}(0.05|0.05)$	367.31	399.48	281.49	6.81
Full estimator				
Mean	399.75	436.96	314.68	6.50
Median	388.07	427.38	304.12	6.40
Standard deviation	91.74	89.93	70.86	0.97
True γ				
Mean	358.57	385.59	279.71	5.82
Median	358.55	384.36	278.54	5.83
Standard deviation	30.23	26.58	28.31	0.29
True $\eta_{\mathbf{p}}^*$				
Mean	392.09	436.92	295.66	6.40
Median	382.37	427.34	289.91	6.33
Standard deviation	84.10	89.92	57.02	0.85
True $\eta_{\mathbf{p}}$				
Mean	394.73	436.94	304.18	6.49
Median	384.88	427.36	298.19	6.42
Standard deviation	84.81	89.92	59.09	0.87

A more detailed view of the performance of the ECQ estimator as well as its components is given by sampling density plots in Figure 1. In the rightmost panels, we indicate the first approximation of ECQ, $Q_{Y|X}^*(p_2|p_1) := Q_Y(p_2\eta_{\mathbf{p}}^*)$, based on the true values of the adjustment factor $\eta_{\mathbf{p}}^*$ and

$(1 - p_2\eta_{\mathbf{p}}^*)$-quantile of the distribution of Y; see (2.5). Note that, apart from being used to explore performance of estimators, distances between $\eta_{\mathbf{p}}$ and $\eta_{\mathbf{p}}^*$ (see panels in the third column), and $Q_{Y|X}(p_2|p_1)$ and $Q_{Y|X}^*(p_2|p_1)$ can also be used to indicate how well the upper tail dependence function approximates the conditional tail probability when p_1 and p_2 are small. Based on the sampling densities, we observe the presence of a small positive bias in the Hill estimator of the tail index γ. This could potentially be remedied by the use of a bias-corrected estimator. The estimates of $\eta_{\mathbf{p}}^*$ tend to have a very small range and any deviations between $\eta_{\mathbf{p}}^*$ and $\eta_{\mathbf{p}}$ result in only minor differences in values of $Q_{Y|X}(p_2|p_1)$ and $Q_{Y|X}^*(p_2|p_1)$. It is interesting to note that, in the case of the Hüsler-Reiss, asymmetric logistic and bivariate t distributions, $\hat{\eta}_{\mathbf{p}}^*$ appears to be more accurate for the exact value $\eta_{\mathbf{p}}$ rather than $\eta_{\mathbf{p}}^*$, although the overall influence of $\hat{\eta}_{\mathbf{p}}^*$ on the estimation of ECQ appears to be quite small.
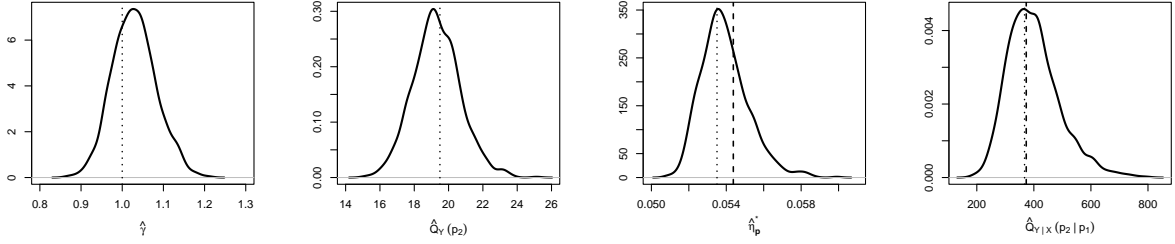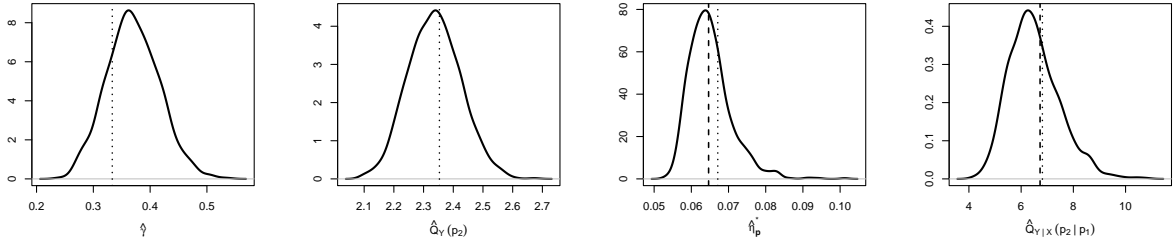

(a) Logistic model

(b) HR model
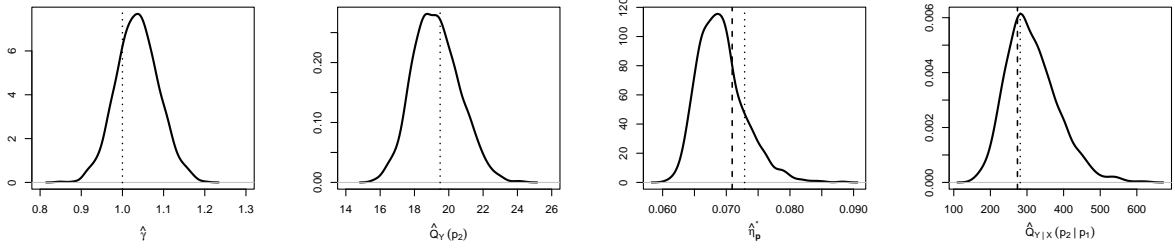
(c) Asymmetric logistic model

(d) Bivariate t model

Figure 1: Sampling densities across 1000 replications of estimates of γ, $Q_Y(p_2)$, $\eta_{\mathbf{p}}^*$ and $Q_{Y|X}(p_2|p_1)$ for $p_1 = p_2 = 0.05$ in the simulation study; see Table 1 for the simulation settings. The dotted vertical lines indicate true values of the quantities being estimated. The dashed vertical lines in the third and fourth column panels indicate the true values of $\eta_{\mathbf{p}}$ and $Q_{Y|X}^*(p_2|p_1)$, respectively.

Finally, we illustrate the asymptotic normality result of Theorem 3.2. Figure 2 displays sampling densities across 1000 replications in the simulation study of the ratio $s_n \left(\frac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)} - 1 \right) / \gamma$ and compares them with the asymptotic standard normal distribution. The sampling densities show a degree of skewness and the coverage rate is somewhat below the nominal level in the asymmetric logistic setting. There also seems to be a bias present for the bivariate t model, where the tail dependence function provides only an approximation to the underlying dependence structure in the tails. However, overall, these results suggest that asymptotic normality could be used as an approximation for inference purposes.
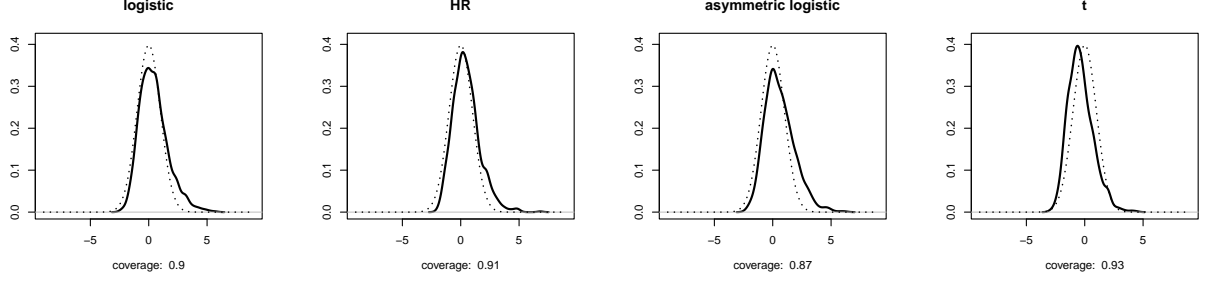
Figure 2: Sampling densities across 1000 replications in the simulation study of the ratio $s_n \left(\dfrac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)} - 1 \right) / \gamma$ (solid lines) along with the standard normal densities (dotted lines) as an illustration of the asymptotic normality result in Theorem 3.2. The empirical coverage rates for the approximate 95% confidence intervals for $Q_{Y|X}(p_2|p_1)$ under the four models are indicated below each plot.

4 Application

The ECQ as defined in (1.1) is closely related to the concept of conditional value-at-risk (CoVaR) in finance. CoVaR at level $1 - p$, denoted $\mathrm{CoVaR}_{Y|X}(p)$, is defined as

$$\mathbb{P}\big(Y \geq \mathrm{CoVaR}_{Y|X}(p) \mid X \geq Q_X(p)\big) = p, \qquad p \in (0,1), \tag{4.1}$$

where $(1-p)$-quantile $Q_X(p)$ is commonly referred to as value-at-risk (VaR) at level $1-p$ for random variable X in financial risk management literature. That is, the two functionals, ECQ and CoVaR, coincide when p_1 is taken to be the same as p_2. It should also be noted that the above definition of CoVaR is different from the original definition of CoVaR introduced by Adrian and Brunnermeier [2016] in that the conditioning distress event is given by the exceedence $\big\{X \geq Q_X(p)\big\}$ rather than equality $\big\{X = Q_X(p)\big\}$. This definition leads to CoVaR being dependence consistent as shown by Mainik and Schaanning [2014]. When defining the CoVaR by conditioning on $\big\{X = Q_X(p)\big\}$, quantile regression techniques are often employed in estimation, see, e.g., Adrian and Brunnermeier [2016] and Leng et al. [2024]. However, quantile regression cannot be used for estimating the CoVaR defined by conditioning on $\big\{X \geq Q_X(p)\big\}$ event. In the subsequent empirical study, we will

19

use the extended definition of CoVaR based on the ECQ functional with two potentially different probability levels p_1 and p_2.

If the focal variable Y is taken as the loss of a proxy for a financial system (e.g., a market index) and the variable in the conditioning event X is the loss of a financial institution, then CoVaR can be used to assess systemic risk, the possibility that an extreme adverse event at an institution level leads to an extreme adverse event for the entire financial system. Regulators rely on measures of system risk to identify what is known as systemically important institutions, which are then subject to stricter regulatory oversight.

In this section, we illustrate how the ECQ estimation methodology of Section 2 can be utilized to produce dynamic CoVaR forecasts using financial time series. In addition to assessing calibration of the resulting forecasts using the unconditional coverage test, we also conduct comparative backtests to compare forecast accuracy among different methods. The considered methods include the ones based on the proposed ECQ estimator with several parametric families for the tail dependence function as well as the one using the empirical estimator of the tail dependence function, a fully parametric method based on AR(1)-GARCH(1,1) filters for the marginals with a bivariate skew-t joint distribution for the residuals[1] and the EVT-based method of Nolde and Zhang [2020]. We refer to the last two methods as "FP" and "EVT-NZ", respectively; both methods were extended to allow for different probability levels in the conditioning and main events to align with the ECQ definition.

We note that both the fully parametric approach using the bivariate skew-t distribution and the method of Nolde and Zhang [2020] which is developed under the framework of multivariate regular variation implicitly assume that the residuals for both institutional and system losses exhibit heavy-tailed behaviour with the same tail index. This assumption may not always be appropriate as illustrated in Figure 3 showing estimates of the tail index for a financial institution and the

---

[1]The fully-parametric method implemented in this section is different from the method in Girardi and Ergün [2013], where a bivariate DCC-GARCH model is utilized to do joint filtering. This is done in order to keep the same filtering procedure across all methods.
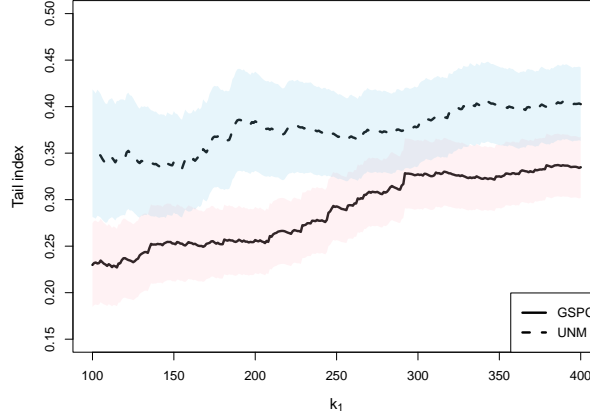
S&P 500 index.

Figure 3: Estimates of the tail index using the Hill estimator in (2.7) as a function of k_1, the number of upper order statistics used in estimation, based on the residuals from the fitted AR(1)-GARCH(1,1) process for the daily losses of a financial institution UnumProvident Corporation (UNM) and the S&P 500 index (GSPC) over the period 2000-2021. The shaded regions indicate asymptotic point-wise 95% confidence bounds.

4.1 Data description

In our application, we consider 15 financial institutions studied in Acharya et al. [2017] with a market capitalization in excess of 5 billion USD as of the end of June 2007, including Aflac Inc. (AFL), American International Group Inc. (AIG), Allstate Corp. (ALL), Bank Of America Corp. (BAC), Citigroup Inc. (C), Comerica Inc. (CMA), Humana Inc. (HUM), JPMorgan Chase & Co. (JPM), Lincoln National Corp. (LNC), Progressive Corp. (PGR), USA Education Inc. (SLM), Travelers Companies Inc. (TRV), UnumProvident Corp. (UNM), Wells Fargo & Co. (WFC) and Washington Mutual Inc. (WM). The S&P 500 index (GSPC) is used as a system proxy. The sample period is from January 1, 2000 to December 30, 2021, consisting of 5535 daily closing price records for each time series. The daily losses (%) were calculated as negative log returns. Figure 4

gives the time series plots of daily losses for one of the institutions (AFL) and the GSPC index, both displaying a typical behavior for financial time series including periods of volatility clustering such as the one during the global financial crisis of 2007–2009.
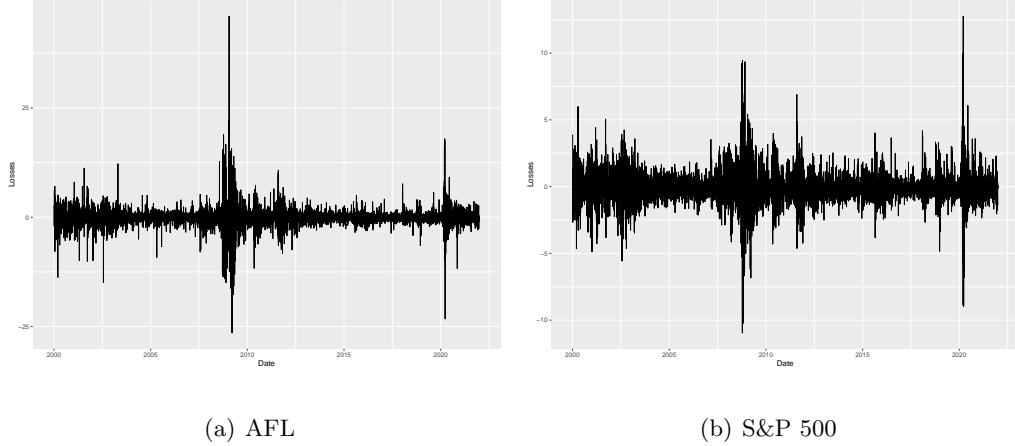

(a) AFL   (b) S&P 500

Figure 4: Time series plots of daily losses (%) for Aflac Inc. (AFL) and the S&P 500 index.

In selection of financial institutions for the data analysis, we considered the length of their available data records as well as compliance with model assumptions for the proposed method. In addition, we chose to include institutions for which tail index estimates differed from the tail index estimate of the S&P 500 index. Tail index estimates are summarized in Table 3.

4.2 CoVaR estimation and forecasting in the dynamic setting

The methodology outlined in Section 2, developed under the premise of i.i.d. observations, is not directly suited to produce dynamic estimates and forecasts of CoVaR for financial time series, known to possess serial dependence and display volatility clustering. One way to address this issue is by combining an ARMA-GARCH model for capturing the evolution of the conditional mean and variance of the underlying stochastic process with an EVT-based static treatment of the i.i.d. innovations; see, e.g. McNeil and Frey [2000]. This leads to a two-stage procedure in which first an ARMA-GARCH process is fitted to the returns (or losses) data, assuming a parametric model for innovations, followed by applying an EVT-based estimation procedure to the sample of realized innovations.

Let $\{X_t^i\}_{t \in \mathbb{N}}$ and $\{X_t^s\}_{t \in \mathbb{N}}$ denote time series of losses for an institution (or company) and a system proxy (market index) adapted to the filtrations $\mathcal{F}^i = \{\mathcal{F}_t^i\}_{t \in \mathbb{N}}$ and $\mathcal{F}^s = \{\mathcal{F}_t^s\}_{t \in \mathbb{N}}$, respectively. To produce dynamic forecasts, we next define conditional versions of risk measures at time t given information in the series up to time $t - 1$. The (conditional) VaR at confidence level $p_1 \in (0, 1)$ for X_t^i given information on the institution's losses up to time $t - 1$, denoted $\text{VaR}_t^i(p_1)$, is defined as the $(1 - p_1)$-quantile of the distribution of X_t^i conditional on $\mathcal{F}_{t-1}^i$:

$$\mathbb{P}\big(X_t^i \geq \text{VaR}_t^i(p_1) \mid \mathcal{F}_{t-1}^i\big) = p_1,$$

and $\text{CoVaR}_t^{s|i}(p_2|p_1)$ is defined as the $(1 - p_2)$-quantile of the conditional loss distribution given information on losses up to time $t - 1$ for both the institution and the system proxy:

$$\mathbb{P}\big(X_t^s \geq \text{CoVaR}_t^{s|i}(p_2|p_1)|X_t^i \geq \text{VaR}_t^i(p_1); \ \mathcal{F}_{t-1}^i, \ \mathcal{F}_{t-1}^s\big) = p_2. \tag{4.2}$$

The details of the two-stage procedure for estimating $\text{CoVaR}_t^{s|i}(p_2|p_1)$ are provided in the Supplementary Material Section S3, with a discussion on its theoretical validity.

4.3 Model diagnostics

Before applying the proposed methodology to the data, we first check appropriateness of several underlying assumptions on the basis of realized innovations from the AR(1)-GARCH(1,1) filter with skew-t distributed innovations. These include independence (lack of serial dependence) and tail dependence.

The sample autocorrelation function (acf) plots and Ljung–Box test validate the assumption of no serial correlation for most of the considered financial institutions with the exception of losses on the Travelers Companies Inc. (TRV) stock and the S&P 500 index. In these two cases, serial dependence can be better captured using an asymmetric GARCH model, such as the exponential GARCH (eGARCH) process. However, the use of the eGARCH filter did not lead to an improvement in the out-of-sample performance of the CoVaR estimators and hence we chose to use the same (symmetric) AR(1)-GARCH(1,1) filter for all time series.

In order to validate the assumption of tail dependence, we estimate the upper tail dependence coefficient, $R(1,1)$ using the nonparametric estimator in (2.9). The resulting estimated values are given in Table 3. For all institutions, the upper tail dependence coefficient estimates are significantly greater than zero (based on a 5000-fold bootstrapping scheme), thus validating tail dependence between losses for each institution and the market index.

Table 3: Estimates of the tail index γ for the 15 considered financial institutions and S&P 500 index (GSPC) and of the upper tail dependence coefficient $R(1,1)$ between the institutional losses and those of the index using the nonparametric estimator in (2.9) with $m = 200$.

Stock	AFL	AIG	ALL	BAC	C	CMA	HUM	JPM
$\hat{\gamma}$	0.353	0.349	0.336	0.308	0.324	0.335	0.356	0.307
$\hat{R}(1,1)$	0.380	0.395	0.295	0.435	0.450	0.410	0.205	0.500
Stock	LNC	PGR	SLM	TRV	UNM	WFC	WM	GSPC
$\hat{\gamma}$	0.314	0.342	0.330	0.345	0.369	0.277	0.359	0.257
$\hat{R}(1,1)$	0.455	0.300	0.305	0.340	0.400	0.390	0.330	$*$

4.4 Results

In this section, we perform a dynamic analysis to assess accuracy of one-day ahead forecasts of (extended) CoVaR at risk level $\mathbf{p} = (0.02, 0.05)$ for the times series described in Section 4.1. The risk level is chosen so as to enable backtesting with sufficiently many observations in the marginal and joint tails. The performance of the estimator in (2.6) under several parametric models for the tail dependence function is assessed via the unconditional coverage test with comparisons of forecast accuracy made using average quantile scores; see Banulescu-Radu et al. [2020] and Fissler and Hoga [2021] for details on backtesting of CoVaR.

We use a rolling window of 3000 data points to estimate model parameters and produce one-day ahead CoVaR forecasts according to the two-stage procedure outlined in Section 4.2 with further details available in the Supplementary Material Section S3. The choice of the size of the estimation

window balances the number of observations available for estimation (note that after conditioning on exceedances over 98% VaR, we expect only about 60 excesses with which to estimate the conditional 95% quantile) with the out-of-sample size of about 2500 observations available for backtesting. To reduce computational time, CoVaR and VaR estimates based on the samples of realized innovations are updated only every 50 observations. Estimation of VaR and CoVaR requires choosing suitable values for sample fractions k_1 and k_2. While it is common to take $k_2 = k_1$ and select a value with the two-step subsample bootstrap algorithm such as in Danielsson et al. [2001], we have found that this procedure leads to very low values of k_1 and k_2 for the considered datasets. As a result, in the present data analysis, we allow the two values to be different and select them separately. We first select a value of k_1 for the tail index estimation using the Hill plot. It suggests that $k_1 = 150$ is a good choice. We found that the performance of the proposed CoVaR estimator is more sensitive to the value of k_2. Thus, we adopt a procedure to select the "best" k_2 by minimizing in-sample average quantile scores.

More specifically, for each rolling window of 3000 data points, we first obtain the realized innovations by applying the AR(1)-GARCH(1,1) filter. We then calculate CoVaR estimates at risk level $\mathbf{p} = (0.02, 0.05)$ over a sequence of k_2 values from 200 to 300. For each CoVaR estimate, an in-sample average quantile score can be evaluated using the classical 1-homogeneous scoring function for $(1 - p_2)$-quantile: $S(r, x) = (p_2 - \mathbb{1}\{x > r\})r + \mathbb{1}\{x > r\}x$ with r denoting the estimate or forecast and x the observation. Finally, we select k_2 as the one whose corresponding CoVaR estimate has the smallest average quantile score. The procedure is repeated for all the 51 windows and five parametric models. We allow k_2 to be different across the 51 windows and the five parametric models. A similar procedure was applied to the EVT-based method in Nolde and Zhang [2020], where k_1 and k_2 are involved in estimating the VaR of institutions.

Note that when calculating the in-sample average quantile score, the VaR estimates for each institution in the conditioning event are also needed, although the proposed CoVaR estimator in (2.6) does not involve explicit computation of the VaR estimates for institutions. Dynamic VaR estimates for institutions are also required for the purpose of out-of-sample backtesting. To keep

consistency across different methods and parametric models, we take the empirical quantiles at level $1 - p_1$ for each rolling window as the VaR estimates used for calculating in-sample average quantile scores and producing dynamic VaR forecasts.

Results of the unconditional coverage tests are summarized in Supplementary Material Section S4. For all institutions other than LNC, the VaR estimates pass the unconditional coverage test at 5% significance level. For most of the institutions, the CoVaR estimates under EVT-NZ method seem to be underestimated, while the fully-parametric estimates seem to be overestimated, leading to several rejections of the unconditional coverage test at level 5% for these two methods. On the other hand, the proposed method with different models for the tail dependence function provides well-calibrated estimates of CoVaR, passing all of the unconditional coverage tests, with the exception of institution HUM modelled using the Hüsler-Reiss tail dependence function.

We would like to remark that while an analysis of the behavior of the CoVaR predictions through time is an important question, it cannot be tackled through simple means such as the Christoffersen test because the notion of transition probability over consecutive days is not applicable.

To make further comparisons across the considered methods, we summarize the average quantile scores for each CoVaR estimator in Table 4. When comparing within the proposed methodology, the asymmetric logistic and t models for the tail dependence function lead to a superior forecasting performance relative to the other three models for most of the financial institutions. Furthermore, when compared with the fully-parametric method and the EVT-based method in Nolde and Zhang [2020], for 12 out of 15 companies, the proposed methodology leads to a better performance in terms of accuracy of CoVaR forecasts.

Figure 5 shows traffic light matrices (see, e.g., Nolde and Ziegel [2017]) for comparative backtests at 10% test confidence level. Comparative backtests assess differences in average quantile scores for each pair of methods. A "proposed" method in a given row is compared with a "reference" method in a given column. A green/red square indicates that the proposed method is significantly better/worse in terms of forecasting accuracy than the reference method. Yellow squares indicate that the score differences are not statistically significant. Figure 5(a) displays the traffic light matrix

Table 4: The average quantile scores of CoVaR forecasts at level $\mathbf{p} = (0.02, 0.05)$. The smallest score for each company is highlighted in bold.

	Log	HR	Alog	t	FP	EVT-NZ
AFL	0.2767	0.2817	0.2654	**0.2645**	0.3087	0.2696
AIG	0.2667	0.2708	**0.2583**	0.2608	0.2903	0.2639
ALL	0.3016	0.3072	0.3000	**0.2980**	0.3191	0.2992
BAC	0.2374	0.2404	**0.2234**	0.2276	0.2374	0.2406
C	0.2412	0.2440	**0.2323**	0.2328	0.2396	0.2401
CMA	0.2562	0.2620	**0.2449**	0.2451	0.2626	0.2489
HUM	**0.2409**	0.2517	0.2473	0.2592	0.2781	0.2592
JPM	0.2325	0.2343	**0.2244**	0.2267	0.2351	0.2362
LNC	0.2319	0.2336	**0.2200**	0.2240	0.2385	0.2262
PGR	0.2673	0.2741	0.2657	**0.2653**	0.2740	0.2687
SLM	0.2366	0.2452	0.2306	0.2350	0.2516	**0.2283**
TRV	0.2786	0.2843	**0.2768**	0.2843	0.2834	0.2826
UNM	0.2383	0.2419	**0.2215**	0.2256	0.2554	0.2337
WFC	0.2365	0.2394	0.2246	0.2306	0.2327	**0.2224**
WM	0.2630	0.2749	0.2526	0.2554	0.2937	**0.2502**

for UnumProvident Corp. (UNM). It shows that the forecasting procedure using the proposed CoVaR estimator with the asymmetric logistic tail dependence function leads to a statistically significant superior forecasting performance compared to the other considered procedures. We note that this result is also supported through assessment of data properties and how they relate to various model assumptions. Figure 3, referred to earlier in the Introduction, gives evidence of potentially different tail indices for the UNM losses and those of the S&P 500 index, which would violate model assumptions for the fully parametric approach and the method in Nolde and Zhang [2020]. Furthermore, carrying out a formal test in Einmahl et al. [2021] results in rejection of the null hypothesis of multivariate regular variation for these data.

While conceptually CoVaR can be backtested in the same way as VaR, conditioning on an

institution's losses being above its VaR estimate or forecast creates a practical difficulty to obtaining conclusive results when performing comparative backtesting due to substantial reduction in the size of the testing data set. For many financial institutions in our application, traffic light matrices tend to contain many yellow cells making comparisons statistically inconclusive but indicating that the proposed method is not worse than the other competing approaches. If one is interested in a method that performs best across all institutions, one possibility is to combine normalized scores. Pooling information in such a way leads to more conclusive results. Figure 5(b) shows the traffic light matrix for comparative backtests based on the normalized average scores combined across all institutions. It confirms that the proposed method with the asymmetric logistic model for the tail dependence function is significantly superior to all of the other methods in terms of its forecasting performance.
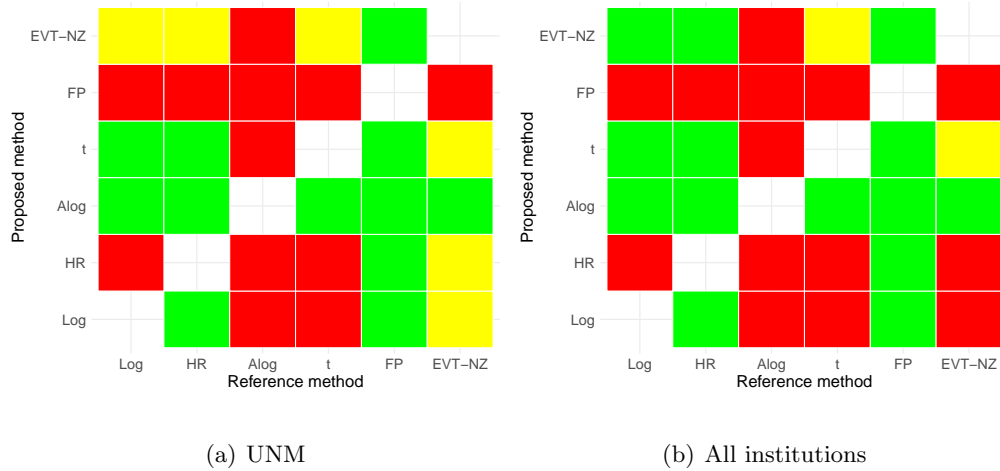

(a) UNM   (b) All institutions

Figure 5: Traffic light matrices for comparative backtests of CoVaR one-day ahead forecasts at risk level $\mathbf{p} = (0.02, 0.05)$ and confidence test level of 10%. The left panel is the result for institution UNM; the right panel is the result using the combined normalized scores across all considered institutions.

5   Conclusion

The paper develops an EVT-based method for estimating extreme conditional quantiles when a related variable is also extreme. Our methodology utilizes the observation that a conditional quantile can be related to the marginal quantile at an adjusted probability level. The adjustment factor is determined by the degree of (tail) dependence between the two underlying variables. A parametric modelling of the tail dependence function addresses the data sparsity issue in the joint tail regions. Expressing the ECQ as a quantile under the marginal distribution allows to then apply nonparametric extrapolation techniques in univariate tail estimation. The asymptotic theory is developed for the proposed ECQ estimator, including consistency and asymptotic normality. Simulation studies illustrate good performance of the estimator and indicate that its bias and variance are dominated by that of the tail index estimator. Using time series data for 15 financial institutions, we find that the proposed method provides a highly competitive alternative to other existing approaches, while allowing for more flexible model assumptions.

We note that parametric modelling of the tail dependence function comes with the challenge of model selection as well as computational complexity. The latter is especially an issue when the dimension of the parameter vector of the selected model is large and the tail dependence function does not have a simple explicit form to carry through the M-estimation. Finding a more flexible yet computationally tractable way to model tail dependence may help to improve the current framework. Another limitation of the proposed methodology is that it can only be applied to situations in which the tail dependence function is non-degenerate, i.e., in the presence of tail dependence. While tail dependence is a reasonable and most relevant assumption in the context of systemic risk, in some situations, for example, in environmental applications, extending the current methodology to also include the case of tail independence will be useful.

Supplementary Material

In the Supplementary Material, Section S1 provides (upper) tail dependence functions for the five parametric models used in simulation studies and the application. Section S2 contains plots of tail dependence function $R(1, \eta)$ as a function of η under several tail dependence models. Section S3 details the two-stage procedure for producing dynamic CoVaR forecasts. Section S4 presents results of the unconditional coverage tests in the application. Section S5 summarizes results of a simulation study under model misspecification.

Acknowledgements

References

V.V. Acharya, L.H. Pedersen, T. Philippon, M. Richardson. Measuring systemic risk. The Review of Financial Studies, 30:2–47, 2017.

T. Adrian and M.K. Brunnermeier. CoVaR. American Economic Review, 106:1705–1741, 2016.

D. Banulescu-Radu, C. Hurlin, J. Leymarie, and O. Scaillet. Backtesting marginal expected shortfall and related systemic risk measures. Management Science, 67:5730–5754, 2020.

J. Cai, 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: 417–442, 2014.

S.G. Coles and J.A. Tawn. Modelling extreme multivariate events. Journal of the Royal Statistical Society: Series B, 53:377–392, 1991.

J. Danielsson, L. de Haan, L. Peng, and C.G. de Vries. Using a bootstrap method to choose the sample fraction in tail index estimation. Journal of Multivariate Analysis, 76:226–248, 2001.

L. de Haan and A. Ferreira. Extreme Value Theory: An Introduction. Springer Science & Business Media, 2006.

J.H. Einmahl, A. Krajina, and J. Segers. A method of moments estimator of tail dependence. Bernoulli, 14:1003–1026, 2008.

J.H. Einmahl, A. Krajina, and J. Segers. An M-estimator for tail dependence in arbitrary dimensions. The Annals of Statistics, 40:1764–1793, 2012.

J.H. Einmahl, A. Kiriliouk, and J. Segers. A continuous updating weighted least squares estimator of tail dependence in high dimensions. Extremes, 21:205–233, 2018.

J.H. Einmahl, F. Yang, and C. Zhou. Testing the multivariate regular variation model. Journal of Business & Economic Statistics, 39:907–919, 2021.

T. Fissler and Y. Hoga. Backtesting systemic risk forecasts using multi-objective elicitability. arXiv preprint arXiv:2104.10673, 2021.

G. Girardi and A.T. Ergün. Systemic risk measurement: Multivariate GARCH estimation of CoVaR. Journal of Banking & Finance, 37:3169–3180, 2013.

B.V. Gnedenko. Sur la distribution limité du terme d'une série aléatoire. Ann. Math., 44:423–453, 1943.

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

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

J. Hüsler and R.D. Reiss. Maxima of normal random vectors: between independence and complete dependence. Statistics & Probability Letters, 7:283–286, 1989.

H. Joe, R.L. Smith, and I. Weissman. Bivariate threshold methods for extremes. Journal of the Royal Statistical Society: Series B, 54:171–183, 1992.

H. Joe. Multivariate Models And Multivariate Dependence Concepts. Chapman and Hall/CRC, 1997.

A.W. Ledford and J.A. Tawn. Statistics for near independence in multivariate extreme values. Biometrika, 83:169–187, 1996.

X. Leng, Y. He, Y. Hou and L. Peng. Asymptotics of CoVaR Inference In Two-Quantile-Regression. Available at SSRN: https://ssrn.com/abstract=4816475 , 2024.

G. Mainik and E. Schaanning. On dependence consistency of CoVaR and some other systemic risk measures. Statistics & Risk Modeling, 31:49–77, 2014.

A.J. McNeil and R. Frey. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance, 7:271–300, 2000.

A.J. McNeil, R. Frey, and P. Embrechts. Quantitative Risk Management: Concepts, Techniques And Tools. Princeton University Press, Princeton, 2005.

N. Nolde and J. Zhang. Conditional extremes in asymmetric financial markets. Journal of Business & Economic Statistics, 38: 201–213, 2020.

N. Nolde and J. Ziegel. Elicitability and backtesting: Perspectives for banking regulation. The Annals of Applied Statistics, 11:1833–1874, 2017.

S.I. Resnick. Extreme Values, Regular Variation and Point Processes. Springer, 1987.

R.L. Smith. Extreme value theory. Handbook of Applicable Mathematics, 7:437–471, 1990.

R.L. Smith. Multivariate threshold methods. In Extreme Value Theory and Applications, pages 225–248. Springer, 1994.

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

Appendix A Proof of Theorems 3.1

Recall the definition of $\eta_{\mathbf{p}}$:

$$\eta_{\mathbf{p}} = \frac{\mathbb{P}\big(Y \geq Q_{Y|X}(p_2|p_1)\big)}{\mathbb{P}\big(Y \geq Q_{Y|X}(p_2|p_1) \mid X \geq Q_X(p_1)\big)} \in (0,1], \qquad p_1, p_2 \in (0,1).$$

We can then rewrite the ratio $\frac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)}$ as

$$
\begin{aligned}
\frac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)} &= \frac{(\eta_{\mathbf{p}}^*)^{-\hat{\gamma}} \, \widehat{Q}_Y(p_2)}{Q_Y(p_2\eta_{\mathbf{p}})} \\
&= \left(\frac{\eta_{\mathbf{p}}^*}{\eta_{\mathbf{p}}}\right)^{-\hat{\gamma}} \times \eta_{\mathbf{p}}^{\gamma-\hat{\gamma}} \times \frac{\widehat{Q}_Y(p_2)}{Q_Y(p_2)} \times \frac{(\eta_{\mathbf{p}})^{-\gamma} Q_Y(p_2)}{Q_Y(p_2\eta_{\mathbf{p}})} \\
&=: I_1 \times I_2 \times I_3 \times I_4.
\end{aligned}
$$

The theorem is proved by showing that, as $n \to \infty$, $I_j \xrightarrow{\mathbb{P}} 1$ for $j = 1,2,3,4$.

Before handling these four terms, the following two lemmas provide some preliminary results regarding the quantities $\eta_{\mathbf{p}}$ and $\eta_{\mathbf{p}}^*$ as well as the estimator $\hat{\eta}_{\mathbf{p}}^*$.

Lemma A.1. *Under the same conditions as in Theorem 3.1, we have that, as $n \to \infty$,*

$$\frac{p_1}{\eta_{\mathbf{p}}^*} \to R_2(1,0;\boldsymbol{\theta}_0) \text{ and } \frac{p_1}{\hat{\eta}_{\mathbf{p}}^*} \xrightarrow{\mathbb{P}} R_2(1,0;\boldsymbol{\theta}_0).$$

Lemma A.2. *Under the same conditions as in Theorem 3.1, we have, as $n \to \infty$,*

$$\frac{\eta_{\mathbf{p}}}{\eta_{\mathbf{p}}^*} \to 1.$$

Proof of Lemma A.1. In order to prove the limit relation regarding $\eta_{\mathbf{p}}^*$, we first show that as $n \to \infty$, $\eta_{\mathbf{p}}^* \to 0$. If otherwise, then there exists a subsequence of integers, $\{n_l\}$ such that $\eta_{\mathbf{p}(n_l)}^* \to c' > 0$ as $l \to \infty$. W.l.o.g., we still use the notation n instead of n_l. Then, as $n \to \infty$, $R\left(1, \eta_{\mathbf{p}}^* \frac{p_2}{p_1}; \boldsymbol{\theta}_0\right) \to R\left(1, \frac{c'}{c}; \boldsymbol{\theta}_0\right) > 0$ which follows from $\frac{p_1}{p_2} \to c$ in Condition C and the fact that $R(1,y;\boldsymbol{\theta}_0)$ is a non-decreasing function in y. However, this contradicts with $R\left(1, \eta_{\mathbf{p}}^* \frac{p_2}{p_1}; \boldsymbol{\theta}_0\right) = p_2 \to 0$ as $n \to \infty$. Hence, we conclude that $\eta_{\mathbf{p}}^* \to 0$ as $n \to \infty$.

Using the mean value theorem, we have that there exists a series of constants $\xi_n \in [0, \eta_{\mathbf{p}}^* \frac{p_2}{p_1}]$ such that

$$p_2 = R\left(1, \eta_{\mathbf{p}}^* \frac{p_2}{p_1}; \boldsymbol{\theta}_0\right) = R(1,0;\boldsymbol{\theta}_0) + \eta_{\mathbf{p}}^* \frac{p_2}{p_1} R_2(1,\xi_n;\boldsymbol{\theta}_0) = \eta_{\mathbf{p}}^* \frac{p_2}{p_1} R_2(1,\xi_n;\boldsymbol{\theta}_0).$$

Hence we get that, as $n \to \infty$,

$$\frac{p_1}{\eta_{\mathbf{p}}^*} = R_2(1, \xi_n; \boldsymbol{\theta}_0) \to R_2(1, 0; \boldsymbol{\theta}_0).$$

Here in the last step, we use the fact that $\xi_n \to 0$ as $n \to \infty$ and $R_2(x, y; \boldsymbol{\theta}_0)$ is a continuous function at $(1, 0; \boldsymbol{\theta}_0)$.

The proof for the limit relation regarding $\hat{\eta}_{\mathbf{p}}^*$ follows similarly by replacing $\boldsymbol{\theta}_0$ with $\hat{\boldsymbol{\theta}}$ and using the fact that $\hat{\boldsymbol{\theta}} \xrightarrow{\mathbb{P}} \boldsymbol{\theta}_0$ as $n \to \infty$. We therefore omit the details. $\square$

Proof of Lemma A.2. We first show that, as $n \to \infty$, $\eta_{\mathbf{p}} \to 0$. If assuming otherwise, there exists a subsequence of integers, $\{n_l\}$ such that $\eta_{\mathbf{p}(n_l)} \to c > 0$ as $l \to \infty$. W.l.o.g., we still use the notation n instead of n_l. Recall the definition of $\eta_{\mathbf{p}}$:

$$\frac{\mathbb{P}(X > Q_X(p_1), Y > Q_Y(p_2 \eta_{\mathbf{p}}))}{p_1} = p_2.$$

By taking $n \to \infty$ on both sides of this equation, and using the assumption $\eta_{\mathbf{p}} \to c' > 0$ together with $\frac{p_1}{p_2} \to c$ as $n \to \infty$, we get that $R(1, \frac{c'}{c}; \boldsymbol{\theta}_0) = 0$, which contradicts Condition F and the fact that $R(1, y; \boldsymbol{\theta}_0)$ is a non-decreasing function in y. Hence, we conclude that, as $n \to \infty$, $\eta_{\mathbf{p}} \to 0$.

Next we show, by contradiction, that

$$\limsup_{n \to \infty} \frac{\eta_{\mathbf{p}}}{\eta_{\mathbf{p}}^*} \le 1.$$

If assuming otherwise, there exists a subsequence of n, $\{n_l\}_{l=1}^{\infty}$ such that as $l \to \infty$, $n_l \to \infty$ and

$$\frac{\eta_{\mathbf{p}(n_l)}}{\eta_{\mathbf{p}(n_l)}^*} \to c' > 1.$$

W.l.o.g., we still use the notation n for the subsequence, and omit it by writing $\mathbf{p} = \mathbf{p}(n)$. Therefore, for any $1 < \tilde{c} < c$, there exists $n_0 = n_0(\tilde{c})$ such that for $n > n_0$, $\frac{\eta_{\mathbf{p}}}{\eta_{\mathbf{p}}^*} > \tilde{c}$.

Note that $\eta_{\mathbf{p}} > \tilde{c} \eta_{\mathbf{p}}^* > \eta_{\mathbf{p}}^*$. By the mean value theorem, we get that for each n, there exists $\xi_n \in \left(\frac{p_2}{p_1} \eta_{\mathbf{p}}^*, \frac{p_2}{p_1} \eta_{\mathbf{p}} \right)$ such that

$$R\left(1, \frac{p_2}{p_1} \eta_{\mathbf{p}}; \boldsymbol{\theta}_0\right) - R\left(1, \frac{p_2}{p_1} \eta_{\mathbf{p}}^*; \boldsymbol{\theta}_0\right) = R_2(1, \xi_n; \boldsymbol{\theta}_0) \frac{p_2}{p_1} (\eta_{\mathbf{p}} - \eta_{\mathbf{p}}^*). \tag{A.1}$$

2

As $n \to \infty$, since both $\eta_{\mathbf{p}}^* \to 0$ and $\eta_{\mathbf{p}} \to 0$ hold, we get $\xi_n \to 0$. Further note that $\eta_{\mathbf{p}} - \eta_{\mathbf{p}}^* > (\tilde{c}-1)\eta_{\mathbf{p}}^*$.

By applying Lemma A.1 and the continuity of $R_2(x, y; \boldsymbol{\theta})$ at $(1, 0; \boldsymbol{\theta}_0)$, we get that

$$\liminf_{n\to\infty} \frac{R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}; \boldsymbol{\theta}_0\right) - p_2}{p_1} = \liminf_{n\to\infty} \frac{R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}; \boldsymbol{\theta}_0\right) - R\left(1, \eta_{\mathbf{p}}^* \frac{p_2}{p_1}; \boldsymbol{\theta}_0\right)}{p_1}$$

$$= \liminf_{n\to\infty} \frac{R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}; \boldsymbol{\theta}_0\right) - R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}^*; \boldsymbol{\theta}_0\right)}{\eta_{\mathbf{p}}^*} \times \frac{\eta_{\mathbf{p}}^*}{p_1}$$

$$\geq R_2(1, 0; \boldsymbol{\theta}_0)\frac{\tilde{c}-1}{c} \times \frac{1}{R_2(1, 0; \boldsymbol{\theta}_0)} = \frac{\tilde{c}-1}{c} > 0.$$

Since Condition E holds with $\tilde{\rho} > 1$, we get that

$$\lim_{n\to\infty} \frac{R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}; \boldsymbol{\theta}_0\right) - p_2}{p_1} = \lim_{n\to\infty} \frac{1}{p_1}\left(R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}; \boldsymbol{\theta}_0\right) - \frac{1}{p_1}\mathbb{P}\left(X > Q_X(p_1), Y > Q_Y\left(p_1 \cdot \frac{p_2}{p_1}\eta_{\mathbf{p}}\right)\right)\right)$$

$$= 0.$$

The two limit relations contradict each other. Therefore, we conclude that

$$\limsup_{n\to\infty} \frac{\eta_{\mathbf{p}}}{\eta_{\mathbf{p}}^*} \leq 1.$$

Similarly, one can show a lower bound for $\frac{\eta_{\mathbf{p}}}{\eta_{\mathbf{p}}^*}$, which completes the proof of the lemma. $\qquad\square$

Now we turn to prove the main theorem by handling the four terms I_j, $j = 1, 2, 3, 4$.

Firstly, we handle I_1. Following the asymptotic property of the Hill estimator (e.g., Theorem 3.2.5 in de Haan and Ferreira [2006]), Condition A and Condition B for k_1 imply that as $n \to \infty$,

$$\sqrt{k_1}(\hat{\gamma} - \gamma) \overset{d}{\to} \mathcal{N}\left(\frac{\lambda_1}{1-\rho}, \gamma^2\right), \tag{A.2}$$

which implies that $\hat{\gamma} \overset{\mathbb{P}}{\to} \gamma$. Together with Lemma A.2 and Lemma A.1, we conclude that $I_1 \overset{\mathbb{P}}{\to} 1$ as $n \to \infty$.

Secondly, we handle I_2. Given the limit relation in (A.2), we only need to show that $\log(\eta_{\mathbf{p}})/\sqrt{k_1} \to 0$ as $n \to \infty$. From Lemma A.1 and Lemma A.2, we get that $\eta_{\mathbf{p}}/p_1 \to 1/R_2(1, 0, \boldsymbol{\theta}_0)$ as $n \to \infty$. Together with the limit relation regarding k_1 in Condition B, we get that $I_2 \overset{\mathbb{P}}{\to} 1$ as $n \to \infty$.

The term I_3 is handled by the asymptotic property of the high quantile estimator; see, e.g. Theorem 4.3.8 in de Haan and Ferreira [2006]. More specifically, under Condition A, Condition B

3

and Condition C, the high quantile estimator in Section 2.3 has the following asymptotic property: as $n \to \infty$,

$$\min\left(\sqrt{k_2}, \frac{\sqrt{k_1}}{\log(k_2/np_2)}\right)\left(\frac{\widehat{Q}_Y(p_2)}{Q_Y(p_2)} - 1\right) = O_P(1).$$

The result follows from the proof of Theorem 4.3.8 in de Haan and Ferreira [2006] with some proper adaptations. A direct consequence is that $I_3 \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$.

Finally, we handle the deterministic term I_4. Notice that $Q_Y(p_2) = U_Y(1/p_2)$ and $Q_Y(p_2\eta_{\mathbf{p}}) = U_Y(1/(p_2\eta_{\mathbf{p}}))$. By applying Condition A with $t = 1/p$ and $x = 1/\eta_{\mathbf{p}}$, we get that

$$\lim_{n \to \infty} \frac{\frac{Q_Y(p_2\eta_{\mathbf{p}})}{Q_Y(p_2)}\eta_{\mathbf{p}}^{\gamma} - 1}{A(1/p)} = -\frac{1}{\rho}.$$

As $n \to \infty$, since $A(1/p) \to 0$ we get that $I_4 \to 1$. $\qquad\square$

Appendix B   Proof of Theorems 3.2

Similar to the proof of consistency, using the notation $Q_Y(p) = U_Y(1/p)$, we write the ratio $\frac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)}$ as

$$
\begin{aligned}
\frac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)} &= \frac{(\hat{\eta}_{\mathbf{p}}^*)^{-\hat{\gamma}}\widehat{Q}_Y(p_2)}{Q_Y(p_2\eta_{\mathbf{p}})} \\
&= \left(\frac{\hat{\eta}_{\mathbf{p}}^*}{\eta_{\mathbf{p}}}\right)^{-\hat{\gamma}} \times \eta_{\mathbf{p}}^{\gamma-\hat{\gamma}} \times \frac{\widehat{Q}_Y(p_2)}{Q_Y(p_2)} \times \frac{(\eta_{\mathbf{p}})^{-\gamma}Q_Y(p_2)}{Q_Y(p_2\eta_{\mathbf{p}})} \\
&= \left(\frac{\hat{\eta}_{\mathbf{p}}^*}{\eta_{\mathbf{p}}}\right)^{-\hat{\gamma}} \times \eta_{\mathbf{p}}^{\gamma-\hat{\gamma}} \times \frac{Y_{n,n-k_2}}{U_Y(n/k_2)} \times \left(\frac{k_2}{np_2}\right)^{\hat{\gamma}-\gamma} \times \frac{U_Y(n/k_2)\left(\frac{k_2}{np_2}\right)^{\gamma}}{U_Y(1/p_2)} \times \frac{(\eta_{\mathbf{p}})^{-\gamma}U_Y(1/p_2)}{U_Y(1/(p_2\eta_{\mathbf{p}}))} \\
&= \left(\frac{\hat{\eta}_{\mathbf{p}}^*}{\eta_{\mathbf{p}}}\right)^{-\hat{\gamma}} \times \exp\left\{\log\frac{k_2}{np_2\eta_{\mathbf{p}}} \cdot (\hat{\gamma}-\gamma)\right\} \times \frac{Y_{n,n-k_2}}{U_Y(n/k_2)} \times \frac{U_Y(n/k_2)\left(\frac{k_2}{np_2\eta_{\mathbf{p}}}\right)^{\gamma}}{U_Y(1/(p_2\eta_{\mathbf{p}}))} \\
&=: I_1 \times I_2 \times I_3 \times I_4.
\end{aligned}
$$

The theorem is proved by handling the four terms separately.

We start with handling I_2. By Condtions A'-C' and the asymptotic normality of the Hill estimator, we get that

$$\frac{\sqrt{k_1}}{\log\frac{k_2}{np_2\eta_{\mathbf{p}}}}(I_2 - 1) \xrightarrow{d} \Gamma,$$

4

where $\Gamma \sim \mathcal{N}\left(\frac{\lambda_1}{1-\rho}, \gamma^2\right)$ is the asymptotic limit for the Hill estimator. By Lemmas A.1 and A.2, we get that as $n \to \infty$,

$$\frac{p_1}{\eta_{\mathbf{P}}} \to R_2(1, 0; \boldsymbol{\theta}_0).$$

Condition F' ensures that $R_2(1, 0; \boldsymbol{\theta}_0) > 0$. Since $p_1 \to 0$ as $n \to \infty$, we obtain that $\frac{\log \eta_{\mathbf{P}}}{\log p_1} \to 1$. This allows us to modify the speed of convergence and obtain that as $n \to \infty$,

$$\frac{\sqrt{k_1}}{\log \frac{k_2}{np_1 p_2}}(I_2 - 1) \xrightarrow{d} \Gamma.$$

Next, we handle I_3. By Conditions A'-C', we have that as $n \to \infty$

$$\sqrt{k_2}\,(I_3 - 1) \xrightarrow{d} B,$$

where $B \sim \mathcal{N}(0, \gamma^2)$. In addition, $\operatorname{cov}(B, \Gamma) \neq 0$ if and only if $k_2/k_1 \to c \in (0, 1)$. In all other cases, $\operatorname{cov}(B, \Gamma) = 0$. Note that, if $k_2/k_1 \to c \in (0, 1)$, $r = \infty$ and $s_n = \frac{\sqrt{k_1}}{\log \frac{k_2}{np_1 p_2}}$.

Next, we handle I_4, which is a deterministic term. Similar to handling the I_4 term in the proof of Theorem 3.1, we get that as $n \to \infty$,

$$I_4 - 1 \sim \frac{1}{\rho} A\left(\frac{n}{k_2}\right).$$

Therefore, we have $\sqrt{k_2}(I_4 - 1) \to \frac{\lambda_2}{\rho}$ as $n \to \infty$.

To handle I_1, we need the following two lemmas. They are improved versions of Lemmas A.1 and A.2, respectively. Their proofs are further postponed to the end of this section.

Lemma B.1. *Under the same conditions as in Theorem 3.2, we have that, as $n \to \infty$,*

$$s_n\left(\frac{p_1}{\eta_{\mathbf{P}}^*} - R_2(1, 0; \boldsymbol{\theta}_0)\right) \to 0;$$

$$s_n\left(\frac{p_1}{\hat{\eta}_{\mathbf{P}}^*} - R_2(1, 0; \hat{\boldsymbol{\theta}})\right) \xrightarrow{\mathbb{P}} 0.$$

Lemma B.2. *Under the same conditions as in Theorem 3.2, we have that, as $n \to \infty$,*

$$s_n\left(\frac{\eta_{\mathbf{P}}}{\eta_{\mathbf{P}}^*} - 1\right) \to 0.$$

5

Finally, we use the two Lemmas to handle I_1. Lemma B.1 implies that

$$s_n \left(\frac{\eta_{\mathbf{p}}^*}{\hat{\eta}_{\mathbf{p}}^*} - \frac{R_2(1,0;\hat{\boldsymbol{\theta}})}{R_2(1,0;\boldsymbol{\theta}_0)} \right) \xrightarrow{\mathbb{P}} 0.$$

Following Einmahl et al. [2012], based on Conditions D' and E', we have that as $n \to \infty$

$$\sqrt{m}\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\|_1 = O_P(1).$$

Since $s_n \leq \sqrt{k_2}$ and $k_2/m \to 0$ (Condition G'), as $n \to \infty$, we get that

$$s_n\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\|_1 = o_P(1).$$

Together with the fact that $R_2(1,0,\boldsymbol{\theta}_0) > 0$ (Condition F') and $R_2(1,0,\boldsymbol{\theta})$ is 1-Lipschitz continuous in the neighborhood of $\boldsymbol{\theta}_0$ (Condition G'), we get that as $n \to \infty$,

$$s_n \left(\frac{R_2(1,0;\hat{\boldsymbol{\theta}})}{R_2(1,0;\boldsymbol{\theta}_0)} - 1 \right) \xrightarrow{\mathbb{P}} 0,$$

which further implies that

$$s_n \left(\frac{\eta_{\mathbf{p}}^*}{\hat{\eta}_{\mathbf{p}}^*} - 1 \right) \xrightarrow{\mathbb{P}} 0.$$

Combining with Lemma B.2, we get that as $n \to \infty$

$$s_n \left(\frac{\eta_{\mathbf{p}}}{\hat{\eta}_{\mathbf{p}}^*} - 1 \right) \xrightarrow{\mathbb{P}} 0,$$

which implies that $s_n(I_1 - 1) \xrightarrow{\mathbb{P}} 0$.

Combining the four terms $I_j, j = 1, 2, 3, 4$, we get that, as $n \to \infty$, if $r \leq 1$, then $s_n = \sqrt{k_2}$ and

$$s_n \left(\frac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)} - 1 \right) \xrightarrow{d} r\Gamma + B + \frac{\lambda_2}{\rho};$$

if $r \geq 1$, then $s_n = \frac{\sqrt{k_1}}{\log \frac{k_2}{np_1p_2}}$ and

$$s_n \left(\frac{\widehat{Q}_{Y|X}(p_2|p_1)}{Q_{Y|X}(p_2|p_1)} - 1 \right) \xrightarrow{d} \Gamma + \frac{1}{r}B + \frac{\lambda_2}{r\rho}.$$

Note that we can regard $\text{cov}(\Gamma, B) = 0$ since the only non-zero case occurs for $r = +\infty$, where the term $\frac{1}{r}B$ vanishes. The bias and variance follow from straightforward calculations.

In the calculation, it is worth noticing that if $r < +\infty$, $k_2/k_1 \to 0$ as $n \to \infty$. In this case, given that λ_1 is finite, we have $\lambda_2 = 0$. In addition, if $r = +\infty$, the term $\frac{\lambda_2}{r\rho}$ vanishes. $\qquad\square$

6

Proof of Lemma B.1. Since $R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}^*; \boldsymbol{\theta}_0\right) = p_2$, by Condition F', we get that as $n \to \infty$,

$$\frac{p_1}{\eta_{\mathbf{p}}^*} - R_2(1, 0; \boldsymbol{\theta}_0) = O((\eta_{\mathbf{p}}^*)^{\breve{\rho}(\boldsymbol{\theta}_0)}) = O(p_1^{\breve{\rho}(\boldsymbol{\theta}_0)}),$$

where the last step follows from Lemma A.1. The first limit relation follows immediately from Condition F'. The second relation follows similarly with replacing $\boldsymbol{\theta}_0$ by $\hat{\boldsymbol{\theta}}_0$. The only difference is that to guarantee that $s_n p_1^{\breve{\rho}(\hat{\boldsymbol{\theta}})} \xrightarrow{\mathbb{P}} 0$, we need the fact that $\hat{\boldsymbol{\theta}}$ is a consistent estimator of $\boldsymbol{\theta}_0$, $\breve{\rho}(\cdot)$ is a continuous function at $\boldsymbol{\theta}_0$ and $s_n p_1^{\breve{\rho}(\boldsymbol{\theta}_0)-\varepsilon} \to 0$ as $n \to \infty$. $\qquad\square$

Proof of Lemma B.2. The proof is similar to that of Lemma A.2, but simpler. Recall Equation (A.1) in the proof of Lemma A.2:

$$R_2(1, \xi_n; \boldsymbol{\theta}_0)\frac{p_2}{p_1}(\eta_{\mathbf{p}}^* - \eta_{\mathbf{p}}) = R\left(1, \frac{p_2}{p_1}\eta_p^*, \boldsymbol{\theta}_0\right) - R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}, \boldsymbol{\theta}_0\right) = p_2 - R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}, \boldsymbol{\theta}_0\right).$$

Therefore, we get

$$s_n\left(\frac{\eta_p^*}{\eta_{\mathbf{p}}} - 1\right) = \frac{1}{R_2(1, \xi_n; \boldsymbol{\theta}_0)} \cdot \frac{p_1}{\eta_{\mathbf{p}}} \cdot s_n \frac{p_2 - R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}, \boldsymbol{\theta}_0\right)}{p_2}$$

$$= I_1 \cdot I_2 \cdot I_3.$$

As $n \to \infty$, $I_1 \to \frac{1}{R_2(1,0;\boldsymbol{\theta}_0)}$ and $I_2 = O(1)$ due to Lemmas A.1 and A.2. Hence we only need to show that $I_3 \to 0$.

Recalling the definition of $\eta_{\mathbf{p}}$ and Condition E', we get that

$$p_2 - R\left(1, \frac{p_2}{p_1}\eta_{\mathbf{p}}, \boldsymbol{\theta}_0\right) = O((\eta_{\mathbf{p}})^{\tilde{\rho}}) = O(p_1^{\tilde{\rho}}),$$

which implies that $I_3 = O(s_n p_1^{\tilde{\rho}-1}) \to 0$ as $n \to \infty$ due to Condition F'. $\qquad\square$

References

L. de Haan and A. Ferreira. Extreme Value Theory: An Introduction. Springer Science & Business Media, 2006.

J.H. Einmahl, A. Krajina, and J. Segers. An M-estimator for tail dependence in arbitrary dimensions. The Annals of Statistics, 40:1764–1793, 2012.

Tail risk in the tail: Estimating high quantiles when a related variable is extreme

Supplementary Material

S1 Parametric models for the tail dependence function

In this section, we provide tail dependence functions for four parametric models considered in simulation studies and the application.

(1) The bivariate logistic distribution function with standard Fréchet margins is given by

$$G(x, y; \theta) = \exp\left\{-(x^{-1/\theta} + y^{-1/\theta})^\theta\right\},$$

where $x, y > 0$ and $\theta \in (0, 1]$. The upper tail dependence function in this case has the form

$$R(x, y; \theta) = x + y - (x^{1/\theta} + y^{1/\theta})^\theta. \tag{S1}$$

(2) The bivariate Hüsler-Reiss distribution function with standard Fréchet margins is

$$G(x, y; \theta) = \exp\left\{-x^{-1}\Phi\left(\theta^{-1} + \frac{\theta}{2}\log(y/x)\right) - y^{-1}\Phi\left(\theta^{-1} + \frac{\theta}{2}\log(x/y)\right)\right\},$$

where $x, y > 0$, $\theta > 0$ and $\Phi(\cdot)$ is the standard normal distribution function. Its tail dependence function is given by

$$R(x, y; \theta) = x + y - x\Phi\left(\theta^{-1} + \frac{\theta}{2}\log(x/y)\right) - y\Phi\left(\theta^{-1} + \frac{\theta}{2}\log(y/x)\right). \tag{S2}$$

(3) The bivariate asymmetric logistic distribution with standard Fréchet margins has distribution function of the form

$$G(x, y; \psi_1, \psi_2, \theta) = \exp\left\{-(1 - \psi_1)/x - (1 - \psi_2)/y - \left((\psi_1/x)^{1/\theta} + (\psi_2/y)^{1/\theta}\right)^\theta\right\},$$

where $x, y > 0$, $\theta \in (0, 1]$ and $\psi_1, \psi_2 \in [0, 1]$. Its tail dependence function is given by

$$R(x, y; \psi_1, \psi_2, \theta) = \psi_1 x + \psi_2 y - \left((x\psi_1)^{1/\theta} + (y\psi_2)^{1/\theta}\right)^\theta. \tag{S3}$$

1

(4) The joint density function of a standard bivariate t distribution with $\nu > 0$ degrees of freedom and correlation parameter $\rho \in (-1, 1)$ is written as

$$f_T(\mathbf{w}; \rho, \nu) = \frac{\Gamma\big((\nu+2)/2\big)}{\sqrt{1-\rho^2}\,\nu\pi\Gamma(\nu/2)}\Big(1 + \frac{1}{\nu}\mathbf{w}^T\Omega^{-1}\mathbf{w}\Big)^{-(\nu+2)/2}, \quad \Omega = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \mathbf{w} \in \mathbb{R}^2.$$

For $\rho \in (0,1)$, its upper tail dependence function is given by

$$R(x, y; \rho, \nu) = xF_T\Big(\sqrt{\frac{\nu+1}{1-\rho^2}}\big(\rho - (y/x)^{-1/\nu}\big); \nu+1\Big) + yF_T\Big(\sqrt{\frac{\nu+1}{1-\rho^2}}\big(\rho - (x/y)^{-1/\nu}\big); \nu+1\Big),$$

(S4)

where $F_T(\cdot; \nu)$ is the distribution function of the standard Student t distribution with ν degrees of freedom. Expression (S4) is founded following Demarta and McNeil [2005], where the lower tail dependence function of the bivariate t distribution is given.

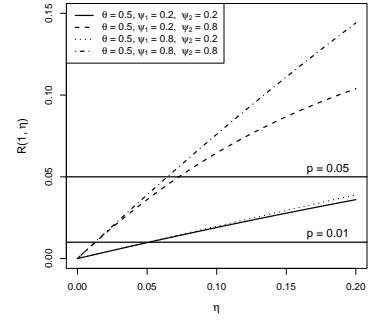## S2   Plots of $R(1, \eta)$ as a function of η

Figure 6 illustrates function $R(1, \eta)$ under the four tail dependence models listed in Section S1 and used in simulation studies and the application. The upper bound $R(1, \eta) \leq \eta$ for $\eta \in (0,1)$ corresponds to the case of complete positive dependence. This implies that when tail dependence is fairly strong, $R(1, \eta)$ is close to the linear function $R(1, \eta) = \eta$. For a given p, η^* in $R(1, \eta^*) = p$ decreases as tail dependence gets stronger.
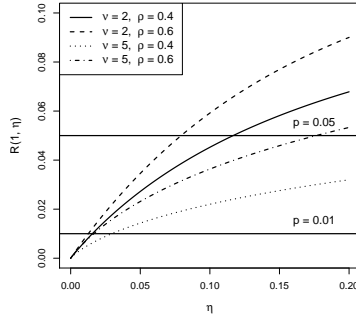
2
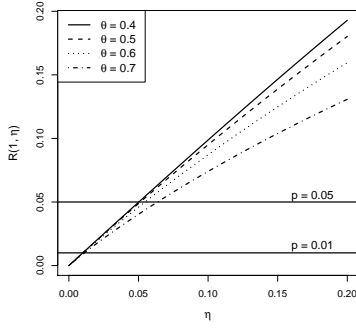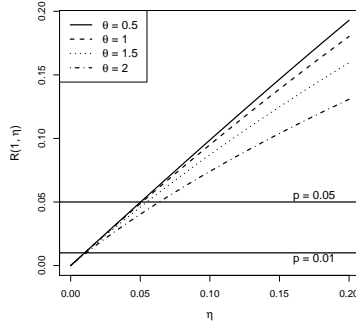
(a) Logistic

(b) Hüsler-Reiss

(c) Asymmetric logistic

(d) Bivariate t

Figure 6: Plots of $R(1, \eta)$ as a function of η under various tail dependence models.

S3 Two-stage procedure for estimating $\text{CoVaR}_t^{s|i}(p_2|p_1)$

To estimate CoVaR dynamically, we need to apply a two-stage procedure, which can be summarized in the steps below.

Step 1 (Univariate GARCH model estimation): Assume that $\{X_t^i\}_{t\in\mathbb{N}}$ and $\{X_t^s\}_{t\in\mathbb{N}}$ each follows an AR(1)-GARCH(1,1) process [Bollerslev, 1986] satisfying the following model equations:

$$X_t^i = \mu_t^i + \sigma_t^i Z_t^i, \quad \mu_t^i = \alpha_0^i + \alpha_1^i X_{t-1}^i, \quad (\sigma_t^i)^2 = \beta_0^i + \beta_1^i(\sigma_{t-1}^i Z_{t-1}^i)^2 + \beta_2^i(\sigma_{t-1}^i)^2,$$

$$X_t^s = \mu_t^s + \sigma_t^s Z_t^s, \quad \mu_t^s = \alpha_0^s + \alpha_1^s X_{t-1}^s, \quad (\sigma_t^s)^2 = \beta_0^s + \beta_1^s(\sigma_{t-1}^s Z_{t-1}^s)^2 + \beta_2^s(\sigma_{t-1}^s)^2,$$

where sequences of innovations $\{Z_t^i\}_{t\in\mathbb{N}}$ and $\{Z_t^s\}_{t\in\mathbb{N}}$ are i.i.d. with zero mean and unit variance. Parameters of the AR(1)-GARCH(1,1) filters are estimated using maximum likelihood assuming a standardized skew-t distribution (Fernández and Steel [1998]) for the innovations. With the estimates of conditional means and volatilities, we can obtain two sequences that could be used as proxies for realized innovations:

$$\big\{\hat{Z}_t^i = (X_t^i - \hat{\mu}_t^i)/\hat{\sigma}_t^i\big\}, \qquad \big\{\hat{Z}_t^s = (X_t^s - \hat{\mu}_t^s)/\hat{\sigma}_t^s\big\}. \tag{S1}$$

Step 2 (Dynamic CoVaR estimation): Based on the time series representation of losses, $\text{CoVaR}_t^{s|i}(p_2|p_1)$ can be expressed as

$$1 - p_2 = \mathbb{P}\big(X_t^s \geq \text{CoVaR}_t^{s|i}(p_2|p_1)|X_t^i \geq \text{VaR}_t^i(p_1); \ \mathcal{F}_{t-1}^i, \ \mathcal{F}_{t-1}^s\big)$$

$$= \mathbb{P}\left(Z_t^s \geq \frac{\text{CoVaR}_t^{s|i}(p_2|p_1) - \mu_t^s}{\sigma_t^s}\Big|Z_t^i \geq \frac{\text{VaR}_t^i(p_1) - \mu_t^i}{\sigma_t^i}; \ \mathcal{F}_{t-1}^i, \ \mathcal{F}_{t-1}^s\right).$$

This suggests first estimating risk measures based on the samples of realized innovations in (S1), treated as i.i.d., and then computing the dynamic forecasts for time t via

$$\widehat{\text{CoVaR}}_t^{s|i}(p_2|p_1) = \hat{\mu}_t^s + \hat{\sigma}_t^s \widehat{\text{CoVaR}}_{Z^s|Z^i}(p_2|p_1), \qquad \widehat{\text{VaR}}_t^i(p_1) = \hat{\mu}_t^i + \hat{\sigma}_t^i \widehat{\text{VaR}}_{Z^i}(p_1). \tag{S2}$$

The two-step procedure is theoretically justified as follows. First, as shown in Hoga [2019] (Proposition 2 in Appendix A), under the assumption that the estimation model is correctly specified, the tail empirical process based on the estimated innovations is arbitrarily close to that based

on the true innovations. The difference is so small that it does not interfere with the asymptotic behavior of the latter process. This result holds for both series. Second, the results above immediately lead to the conclusion that the bivariate tail empirical process based on the two estimated innovations is arbitrarily close to that based on the two true innovations, because the difference is bounded by the sum of the differences in the two marginals. This is an obvious result due to the fact that two indicators of a joint event based on the estimated and true innovation can differ only if at least one pair of marginal indicators differs. In other words, we can show that the bivariate tail empirical process based on the two estimated innovations share the same asymptotic behavior as that based on two true innovations, which has a bivariate Gaussian process as the limit. Such a result is the starting point for proving the asymptotic behavior of the moment estimator as in Einmahl et al. [2012]. Therefore, in the third step, following the lines of the proof in Einmahl et al. [2012], we obtain the same asymptotic result for the moment estimator with the same speed of convergence. Fourth, we can get the asymptotic behavior of all other components needed in the proof of Theorem 3.2 following from the marginal tail empirical process results. Finally, combining all components and again following the proof of Theorem 3.2, one can show the asymptotic normality of the estimator for the time varying CoVaR.

This strategy of proof follows similar ideas for other estimators in the literature. For instance, Hoga [2022] provides an extension to situations where filtering is based on a general location-scale model. See also Girard et al. [2021] for results on residual-based extreme value estimators in heavy-tailed regression models.

We note that in Step 1 above, if there is evidence of time changing correlation structure in the data, an alternative is to use a bivariate GARCH filter as was previously done in Girardi and Ergün [2013] and Nolde and Zhang [2020]. For the data considered here, filtering out correlation led to weaker tail dependence potentially invalidating the assumption of tail dependence (for example, see Figure 7). We, therefore, chose to apply the GARCH filters only marginally.
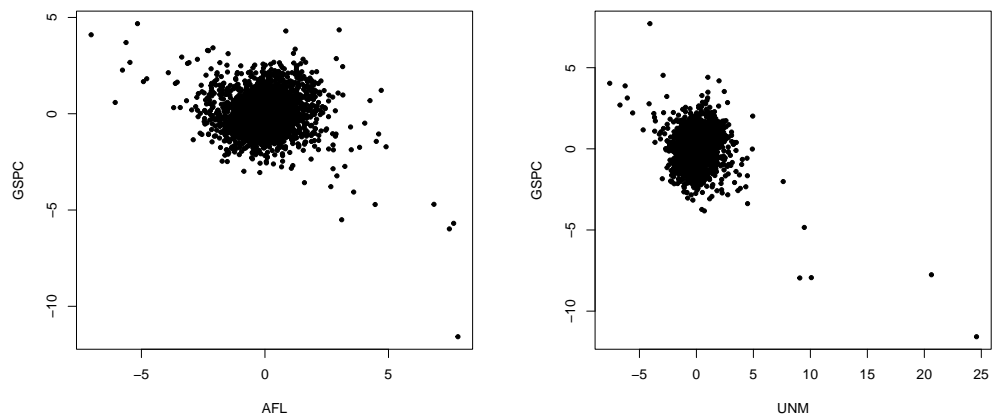
Figure 7: Scatterplots of standardized innovation vectors from GARCH-DCC model for institutions AFLAC INC (AFL) and UNUMPROVIDENT CORP (UNM) with the S&P 500 index (GSPC) as a system proxy.

S4 Unconditional coverage test

Table 5: Unconditional coverage tests for VaR of institutions and CoVaR at level $\mathbf{p} = (0.02, 0.05)$. Tail dependence models include logistic (Log), Hüsler-Reiss (HR), bilogistic (Bilog), asymmetric logistic (Alog) and that of the bivariate t distribution (t); see Supplementary Appendix S1 for model specifications. E_n/e_n denotes the observed/nominal number of exceedances of the VaR estimate, and E_n^b/e_n^b is the observed/nominal number of joint exceedances of VaR and CoVaR estimates.

		AFL	AIG	ALL	BAC	C	CMA	HUM	JPM	LNC	PGR	SLM	TRV	UNM	WFC
VaR	E_n	56	43	57	42	55	62	46	51	66	58	43	56	57	58
	e_n	50.68	50.68	50.68	50.68	50.68	50.68	50.68	50.68	50.68	50.68	50.68	50.68	50.68	50.68
	p-value	0.4578	0.2633	0.3792	0.2045	0.5453	0.1205	0.5	0.9638	**0.0377**	0.3098	0.2633	0.4578	0.3792	0.3098
Log	E_n^b	2	2	3	1	2	2	3	2	2	3	1	3	1	2
	e_n^b	2.8	2.15	2.85	2.1	2.75	3.1	2.3	2.55	3.3	2.9	2.15	2.8	2.85	2.9
	p-value	0.606	0.9155	0.928	0.3877	0.6265	0.4942	0.6503	0.7139	0.4297	0.9522	0.3708	0.9035	0.1965	0.5666
HR	E_n^b	2	2	3	1	2	2	0	2	2	3	1	3	1	2
	e_n^b	2.8	2.15	2.85	2.1	2.75	3.1	2.3	2.55	3.3	2.9	2.15	2.8	2.85	2.9
	p-value	0.606	0.9155	0.928	0.3877	0.6265	0.4942	**NA**	0.7139	0.4297	0.9522	0.3708	0.9035	0.1965	0.5666
Bilog	E_n^b	2	2	3	1	2	2	3	2	2	3	1	3	1	2
	e_n^b	2.8	2.15	2.85	2.1	2.75	3.1	2.3	2.55	3.3	2.9	2.15	2.8	2.85	2.9
	p-value	0.606	0.9155	0.928	0.3877	0.6265	0.4942	0.6503	0.7139	0.4297	0.9522	0.3708	0.9035	0.1965	0.5666
Alog	E_n^b	3	3	4	2	3	4	3	3	2	3	3	4	2	2
	e_n^b	2.8	2.15	2.85	2.1	2.75	3.1	2.3	2.55	3.3	2.9	2.15	2.8	2.85	2.9
	p-value	0.9035	0.5736	0.50890	0.9431	0.8788	0.615	0.6503	0.7782	0.4297	0.9522	0.5736	0.4881	0.586	0.5666
t	E_n^b	3	4	4	1	2	3	3	2	2	3	3	4	1	2
	e_n^b	2.8	2.15	2.85	2.1	2.75	3.1	2.3	2.55	3.3	2.9	2.15	2.8	2.85	2.9
	p-value	0.9035	0.245	0.5089	0.3877	0.6265	0.9533	0.6503	0.7139	0.4297	0.9522	0.5736	0.4881	0.1965	0.5666
FP	E_n^b	1	1	2	1	2	2	0	2	2	3	0	3	0	2
	e_n^b	2.8	2.15	2.85	2.1	2.75	3.1	2.3	2.55	3.3	2.9	2.15	2.8	2.85	2.9
	p-value	0.2058	0.3708	0.5860	0.3877	0.6265	0.4942	**NA**	0.7139	0.4297	0.9522	**NA**	0.9035	**NA**	0.5666
EVT-NZ	E_n^b	4	4	7	6	6	6	4	7	7	4	2	5	6	5
	e_n^b	2.8	2.15	2.85	2.1	2.75	3.1	2.3	2.55	3.3	2.9	2.15	2.8	2.85	2.9
	p-value	0.4881	0.245	**0.0318**	**0.0227**	0.0798	0.1319	0.2956	**0.0174**	0.0672	0.5298	0.9155	0.2221	0.0931	0.2491

S5 Simulation studies under model misspecification

In this section, we report results on the performance of the proposed ECQ estimator under the situation when the fitted model for the tail dependence function is misspecified. We conducted simulation studies using the settings of Section 3.3, but fit just the asymmetric logistic model for data generated using the other three tail dependence models. As summary statistics in Table 6 show, interestingly, there is both a slight lower empirical bias and a lower standard deviation when data come from the logistic and HR distributions compared to the original simulation study in which the correctly specified models were fitted. For the bivariate t distribution, there is a minor increase in bias and variance but overall results are fairly similar to those under the correct model specification. In practice, the choice of the parametric family for the tail dependence function can be based on results of comparative backtests.
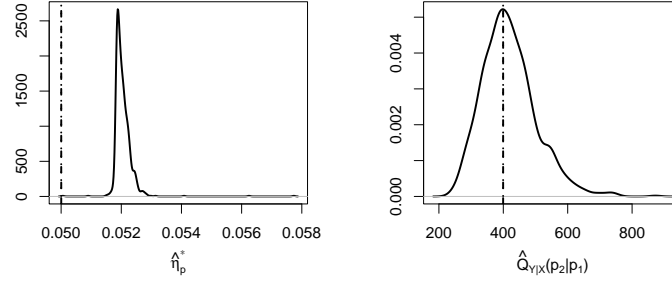
Table 6: Summary statistics of ECQ estimates at level $\mathbf{p} = (0.05, 0.05)$ for simulation studies under model misspecification. The first row gives the true value of the ECQ under each model.

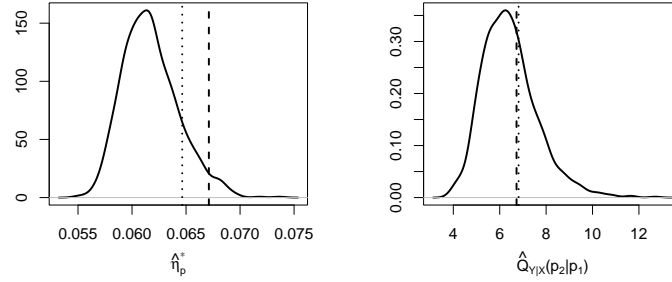|	logistic	HR	bivariate t
$Q_{Y|X}(0.05|0.05)$	367.31	399.48	6.81
Mean	389.01	418.78	6.44
Median	377.18	409.95	6.30
Standard deviation	89.34	84.43	1.20

(a) Logistic model

8
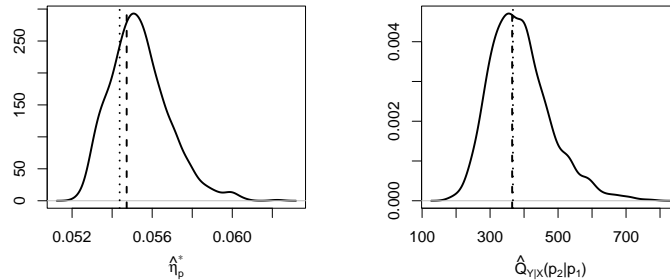
(b) HR model

(c) Bivariate t model

Figure 8: The sampling densities of estimates of γ, η_p^*, $Q_Y(p_2)$ and $Q_{Y|X}(p_2|p_1)$ for $p_1 = p_2 = 0.05$ in the simulation study under model misspecification. The dotted vertical lines indicate true values of the quantities being estimated. The dashed vertical lines indicate the true values of $\eta_{\mathbf{p}}$ and $Q_{Y|X}^*(p_2|p_1)$, respectively.

References

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31:307–327, 1986.

S. Demarta and A. J. McNeil. The t copula and related copulas. International Statistical Review, 73:111–129, 2005.

C. Fernández and M. Steel. On Bayesian modeling of fat tails and skewness. <u>Journal of the American Statistical Association</u>, 93:359–371, 1998.

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

G. Girardi and A.T. Ergün. Systemic risk measurement: Multivariate GARCH estimation of CoVaR. <u>Journal of Banking & Finance</u>, 37:3169–3180, 2013.

Y. Hoga. Confidence intervals for conditional tail risk measures in ARMA–GARCH models. <u>Journal of Business & Economic Statistics</u>, 37(4): 613–624, 2019.

Y. Hoga. Limit theory for forecasts of extreme distortion risk measures and expectiles. <u>Journal of Financial Econometrics</u>, 20(1): 18–44, 2022.

N. Nolde and J. Zhang. Conditional extremes in asymmetric financial markets. <u>Journal of Business & Economic Statistics</u>, 38: 201–213, 2020.