TailCoR

Lorenzo Ricci[1] and David Veredas[2]

First draft: May 2012

This version: May 2015

Abstract

Economic and financial crises are characterized by tail events. When they occur, tail correlations emerge, which have linear and non-linear origins. The former are due to the Pearson correlations, while the strength of the latter depends on the heavyness of the tails. We introduce TailCoR, a new metric for tail correlations that disentangles straightforwardly the linear and non-linear correlations. TailCoR is simple to compute, no optimizations are needed, and it performs well in small samples. When applied to a panel of eight major US banks, TailCoR increases during the financial crisis because of a surge in both the linear and non-linear correlations. The end of 2012 also shows an increase of TailCoR, which is solely driven by the non-linearity, reflecting the risks of tail events and their spillovers associated with the European sovereign debt crisis.

Keywords: Tail correlation, tail risk, quantile, ellipticity, crises.

JEL classification: C32, C51, G01.

[1]ECARES, Solvay Brussels School of Economics and Management, Université libre de Bruxelles (ULB).

[2]ECARES, Solvay Brussels School of Economics and Management, Université libre de Bruxelles (ULB). Corresponding address: David Veredas, ECARES, Solvay Brussels School of Economics and Management, Université libre de Bruxelles, 50 Av F.D. Roosevelt CP114/04, B1050 Brussels, Belgium. Phone: +3226504218. Web: www.ecares.org/veredas.html. Email: david.veredas@ulb.ac.be.

1 Introduction

The 2007-2010 financial and the 2009-2012 European sovereign debt crises highlighted the importance of tail – or rare – events in asset prices. These events have different natures, such as corporate and Government defaults, stock market crashes, or political news, to name a few. When they occur, their effect is spread over the economic and financial systems, creating tail correlations that have linear and non-linear origins. Indeed, tail correlations may happen because asset prices are either linearly correlated (i.e. the Pearson correlations are different from zero) and/or non-linearly correlated, in the sense that asset prices are dependent at the extremes.[1]

The analysis of tail correlations and their linear and non-linear contributions has numerous financial applications, as shown in Poon et al. (2004). For instance, a portfolio of securities that are non-linear risk independent has thinner tails than a portfolio with securities that are non-linear risk dependent. Therefore, the investment decision does not only depend on the investor's risk appetite, but also on her/his preferences for linear and non-linear risks. Disentangling them allows for a more effective risk reduction.

Several measures of comovements have been proposed to assess the degree of tail dependence. The tail dependence coefficients (also called extremal dependence structures and co-excedaance probabilities) steaming from extreme value theory are the most used. These coefficients rely on two building blocks. The first is the assumption that tails (either of the marginal and/or the joint distribution) asymptotically decay according to a power law, and the second is the extreme value copula (also called stable tail dependence function).[2] Broadly speaking, estimation of extreme value copula is divided in two approaches: parametric and semi-parametric. The former is based on parametric copula functions, the logistic and copula-t being commonly used (see, for example, Longin and Solnik (2001) for an application to extreme correlation of international equity markets, and Chollete et al. (2011) for its use in investment diversification across major stock market indexes). The latter is based on higher order statistics (see Embrechts et al. (2000) for a theoretical treatment, and Hartmann et al. (2004) for

[1] Other forms of non-linearity may occur but we do not consider them in this article. They can however be considered within the theory of TailCoR.

[2] See Joe (1997), Embrechts et al. (2000), Beirlant et al. (2004) and McNeil et al. (2005).

an application to asset market linkages in crisis periods).[3] As pointed out in Straetmans et al. (2008), estimation of a tail copula has the weakness that it presupposes tail dependence, besides that the parametric copula rests on parametric assumptions on the unknown dependence function. An alternative, proposed by Ledford and Tawn (1996), consists of testing for tail dependence through the tail index of an auxiliary variable (the cross-sectional minimum of the random vector). Hartmann et al. (2005) takes this avenue for studying the banking system. Poon et al. (2004) and Straetmans et al. (2008) also use Ledford and Tawn (1996) for exploring the tail dependence structure among risky asset returns and US sectoral stock indexes around 9/11 respectively. The use of multivariate extreme value theory has, however, shortcomings. The main one is that the theory builds upon asymptotic results on the tails. Since in practice we estimate using a cut-off point of the tail, the asymptotic results are an approximation. Moreover, precise estimators require large sample sizes.

We introduce TailCoR, a new measure for tail correlations that can be implemented under general and mild assumptions. TailCoR has numerous and useful features: i) it is not an asymptotic (on the tail) dependence measure and hence it does not rely on assumptions of the asymptotic behavior of the tail, ii) it does not depend on specific distributional assumptions, iii) it is simple and no optimizations are needed, iv) it can be computed for tails that are fatter, equal or thinner than Gaussian, and v) it performs well in small samples. Moreover, under the elliptical family of distributions (i.e. the probability contours are ellipsoids) TailCoR disentangles easily between the linear and non-linear correlations.

We also show an extension that differentiates between the downside and upside risks. We call it downside- and upside-TailCoR. It is often the case, like in risk management, where the interest lies in the tail of one region of the distribution, which leads to TailCoR when the probability contours are asymmetric. To cope with this possibility we explain how to disentangle TailCoR in the linear and non-linear correlations under the normal mean-variance mixture distribution.

The empirical illustration is on daily returns of eight major US banks for the period 2003-

[3]Longin and Solnik (2001) exploit, in the bivariate case, the explicit and simple relation between the tail dependence coefficient of the Gumbel copula and the linear correlation coefficient for extreme events, also known as exceedance correlation. In a similar vein, Cizeau et al. (2001) introduce the quantile correlation, i.e. the sample correlation between observations that are contained in a ball around a given joint quantile.

2012. For comparison purposes we compute the upper and lower exceedance correlations, and the parametric and non-parametric tail dependence coefficients (the former with a copula-t and the latter as in Straetmans et al. (2008)). Results for the overall sample show that both the linear and non-linear contributions matter. To study the dynamic behavior we use windows of 6 months. The gains of TailCoR in relatively small samples become apparent. The exceedance correlations could not be computed because of lack of extreme observations. The copula-t tail dependence coefficient shows a great deal of variability across banks from one period to another, and the non-parametric counterpart is almost constant across banks and periods, varying within a tight interval. TailCoR is more stable than the coputal-t and with a dynamic evolution in line with the financial and economic events that happened during the sample. Indeed, during the financial crisis TailCoR increased substantially due a surge in the linear and non-linear contributions. However during the second half of 2012, in the course of the utmost economic and financial troubles in Europe, the increase in TailCoR is solely due to the non-linear contribution. This reflects the increasing concerns of tail events emerging from the Eurozone – namely the fears of an exit of Greece from the euro, the political turmoil in Italy, and the financial bailout in Spain.

The remaining sections are laid out as follows. Section 2 introduces the notation, assumptions, definition, and representations of TailCoR. It also shows a calibration exercise and the asymptotic properties of the estimators. Section 3 covers a brief Monte Carlo study. The extension to asymmetry is touched upon in Section 4. The illustration to the eight US banks is presented in Section 5. Section 6 concludes and lengthy tables and the proofs are relegated to the appendixes.

2 TailCoR

2.1 Definition

Let $\mathbf{X}_t$ $t = 1, \ldots, T$ be a random vector of size N at time t satisfying the following assumption

G1 (a) The random process $\{\mathbf{X}_t\}$ is a strongly stationarity sequence of random vectors, (b) the distribution of $\mathbf{X}_t$ is unimodal, and (c) $\mathbf{X}_t$ is S-mixing.

Assumption **G1**(a) is standard in time series analysis, **G1**(b) is due to Sherman (1955), and **G1**(c) specifies the time dependence of $\mathbf{X}_t$. Unimodality rules out distributions with several modes but not asymmetry. Assuming a mixing condition instead of a particular type of dynamic model makes TailCoR applicable to a wide array of processes. The conditions for S-mixing, introduced by Berkes et al. (2009), apply to a large number of processes used in economics and finance, including GARCH models and its extensions, linear processes (like ARMA models), and stochastic volatility among others. The notation **G** stands for General, as opposed to **E** and **A** for ellipticity and asymmetry that are used below.

TailCoR is based on the following simple idea, shown in Figure 1. If two random variables X_j and X_k (properly standardized) are positively related (either linear and/or non-linearly), most of the times the pairs of observations (depicted with circles) have the same sign, in the sense that most of them concentrate in the north-east and south-west quadrants of the scatter plot. Now, consider the ϕ-degree line that crosses these quadrants (we illustrate the figures with $\phi = \pi/4$, or the 45-degree line), and project all the pairs on this line, producing a new random variable $Z^{(j\,k)}$, depicted with squares.[4] Since the two random variables are positively related, the projected dots –that are sitting on the ϕ-degree line– are dispersed all over the line.[5] The extend of the dispersion depends on the strength of the relation between X_j and X_k. If weak, the cloud of dots is concentrated around the origin and does not have a well defined direction. The dispersion is therefore small. By contrast, if the relation is strong, the cloud is stretched around the ϕ-degree line, and the projected dots are very dispersed.

[FIGURE 1 ABOUT HERE]

TailCoR is equal –up to a normalization– to the difference between the upper and lower tail quantiles of $Z^{(j\,k)}$. The *tail interquantile range* can be large because of two reasons. First, if X_j and X_k are highly linearly correlated, then the dots and the corresponding squares are close to each other (panel (a) of Figure 1). Second, if the correlation between X_j and X_k only happens on the tails, while the observations around the origin form a cloud with undefined

[4]Because of representation purposes we show the projection only for the observations that are far from the origin, but the reader should keep in mind that the projection is done for all the observations.

[5]In the case of negative relation, the dots mostly concentrate in the north-west and south-east quadrants, and the projection is on the corresponding ϕ-degree line, as explained in detail later.

direction (panel b). These two situations are not mutually exclusive and both may happen, which is actually the most likely case in practice (panel c). In either case TailCoR is large, in a sense to be precisely defined below. But first we put this intuition at work.

Let $X_{j\,t}$ be the jth element of the random vector $\mathbf{X}_t$. Denote by Q_j^τ its τth quantile for $0 < \tau < 1$, and let $\mathrm{IQR}_j^\tau = \mathrm{Q}_j^\tau - \mathrm{Q}_j^{1-\tau}$ be the τth interquantile range. A typical value of τ is 0.75. Let $Y_{j\,t}$ be the standardized version of $X_{j\,t}$:

$$Y_{j\,t} = \frac{X_{j\,t} - \mathrm{Q}_j^{0.50}}{\mathrm{IQR}_j^\tau}. \tag{1}$$

Likewise for $Y_{k\,t}$. In our context of heavy tails the median and the interquantile range are used for the standardization. The mean of $Y_{j\,t}$ is not necessarily zero and its variance is not one – if they exist. This is not an issue since the aim of (1) is to have the pair $(Y_{j\,t}, Y_{k\,t})$ centered around the same number and with the same scale. Alternatively, $X_{j\,t}$ and $X_{k\,t}$ could be standardized with the marginal cumulative distribution functions, i.e. $Y_{j\,t}$ and $Y_{k\,t}$ would be the probability integral transforms, which are distributed uniformly on $(0, 1)$. This is advantageous if there are marginal dependencies beyond the location and the scale. Otherwise, as it is in the case of ellipticity discussed below, the standardization (1) is enough.

As (1) is based on marginal quantiles, we need the following technical assumption

G2 (a) For $0 < \tau < 1$, the cumulative distribution function of $X_{j\,t}$ at point $x_{j\,t}$, denoted by $F(x_{j\,t})$, is continuous in some neighborhood of Q_j^τ. (b) The probability density function at $x_{j\,t}$, denoted by $f(x_{j\,t})$, is such that $0 < f(\mathrm{Q}_j^\tau) < \infty$. Likewise for $X_{k\,t}$.

By standard trigonometric arguments, the projection of $(Y_{j\,t}, Y_{k\,t})$ on the ϕ-degree line is

$$Z_t^{(j\,k)} = Y_{j\,t} \cos\phi + Y_{k\,t} \sin\phi, \tag{2}$$

and the tail interquantile range of $Z_t^{(j\,k)}$ is

$$\mathrm{IQR}^{(j\,k)\,\xi} = \mathrm{Q}^{(j\,k)\,\xi} - \mathrm{Q}^{(j\,k)\,1-\xi},$$

where $\mathrm{Q}^{(j\,k)\,\xi}$ is the ξth quantile of $Z_t^{(j\,k)}$, and ξ is larger than τ, typically beyond 0.90. The larger ξ is, the further we explore the tails. Equipped with $\mathrm{IQR}^{(j\,k)\,\xi}$, we define TailCoR as follows.

Definition 1 Under **G1-G2**, TailCoR between $X_{j\,t}$ and $X_{k\,t}$ is

$$\mathrm{TailCoR}^{(j\,k)\,\xi} := s_g(\xi,\tau)\mathrm{IQR}^{(j\,k)\,\xi}, \tag{3}$$

where $s_g(\xi,\tau)$ is a normalization such that under independence $\mathrm{TailCoR}^{(j\,k)\,\xi} = 1$, the reference value.

Four remarks are in order. First, $s_g(\xi,\tau)$ is the inverse of $\mathrm{IQR}^{(j\,k)\,\xi}$ under Gaussianity and linear uncorrelation (and hence independence). A table with values of $s_g(\xi,\tau)$ for a grid of reasonable values for τ and ξ is found in Appendix T. Interpolation can be used for values of ξ and τ that are not in the table or, alternatively, a simple function can be programmed to compute exactly $s_g(\xi,\tau)$ for any value of τ and ξ. The steps for programming it are also given in Appendix T.

Second, the meaning of correlation in TailCoR is not the traditional one, as TailCoR is not bounded between -1 and 1 nor centered at zero. In the next sub-section we deal with this issue in two ways. We first examine the bounds more in detail, and then we propose an alternative representation of TailCoR that is bounded between -1 and 1, and equals 0 under independence.

Third, TailCoR is not an asymptotic (in the tail) dependence measure. Instead, the dependence between $X_{j\,t}$ and $X_{k\,t}$ can be computed for any $\xi < 1$. Moreover, since it does not rely on assumptions of the asymptotic behavior of the tail, TailCoR can be used for distributions with tails that are Pareto, exponential, or even with finite end points.

Last, the angle ϕ has to be chosen. An optimality criteria would be to choose the angle that maximizes the variability of $Z_t^{(j\,k)}$, as it is done in the next sub-section.

2.2 Disentangling the linear and non-linear correlations

Ellipticity and assumptions

Though definition (3) is simple and intuitive, it does not allow to understand how much of TailCoR is due to the linear and non-linear contributions. This disentangling is however possible if we specify further structure on the probability law of $\mathbf{X}_t$, namely ellipticity, leading to a new set of assumptions that we denote by **E**.

E1 The unconditional distribution of $\mathbf{X}_t$ belongs to the elliptical family, given by the stochastic representation $\mathbf{X}_t =_d \boldsymbol{\mu} + \mathcal{R}_{\alpha\, t}\boldsymbol{\Lambda}\mathbf{U}_t$.

The $N \times 1$ random vector $\mathbf{U}_t$ under **E1** is uniformly distributed on the unit sphere. The scaling matrix $\boldsymbol{\Lambda}$ produces the ellipticity and is such that $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}'$, a positive definite symmetric dispersion matrix – often called the shape matrix with generic $(j\, k)$ element $\sigma_{j\, k}$. The non-negative and continuous random variable $\mathcal{R}_{\alpha\, t}$ generates the tail thickness through the shape parameter α, and is stochastically independent of $\mathbf{U}_t$. From now on we will denote α as the tail index, in the sense that this parameter explains the decay of the tails (the smaller α the thicker the tails), but not necessarily according to a power law.[6] The vector $\boldsymbol{\mu}$ re-allocates the center of the distribution. Let $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\Sigma}, \alpha) \in \boldsymbol{\Theta}$ denote the vector of unknown parameters satisfying the following standard assumption.

E2 (a) The parameter space $\boldsymbol{\Theta}$ is a non-empty and compact set on $\mathbb{R}^{N+\frac{N(N+1)}{2}+1}$. (b) The true parameter value $\boldsymbol{\theta}_0$ belongs to the interior of $\boldsymbol{\Theta}$.

Note that assumption **E1** is about a family of distributions, i.e. $\mathbf{X}_t$ is assumed to belong to that family but no specific distributional assumption is made. The elliptical family is commonly used as it nests, among others, the Gaussian, Student-t, elliptical stable (ES henceforth), Cauchy, Laplace, and Kotz probability laws.[7] For a given vector of locations and a dispersion matrix, the difference between two elliptical distributions is the tail index α, which plays a central role all over the remaining of the article. Another feature of the elliptical

[6] The random variable $\mathcal{R}_{\alpha\, t}$ may not only depend on the tail index but in a vector of shape parameters. We do not consider the latter since most of the elliptical distributions used in practice only depend on α (see examples after assumption **E2**). The extension is however straightforward.

[7] See Hashorva (2008) and Hashorva (2010) for tail theory within the elliptical family.

family is its closeness to location and scale shifts, which implies that $Z_t^{(j\,k)}$ is elliptical with the same tail index α.

We also need existence of the mean and the variance-covariance matrix, which is ensured by the following assumption.

E3 The unconditional moments up to order 2 are finite, i.e. $E(\mathbf{X}_t^p) < \infty,$ for $p \leq 2$.

This assumption may be relaxed, as explained in the estimation section. Next, we substitute assumptions **G1** and **G2** for

E4 (a) The random process $\{\mathbf{X}_t\}$ is weakly stationarity sequence of random vectors, and (b) $\mathbf{X}_t$ is S-mixing.

E5 For $0 < \tau < 1$, the cumulative distribution function $P(\mathcal{R}_{\alpha\,t} \leq r)$ has a bounded, continuous and positive derivative.

Assumption **E5** replaces **G2** since it ensures that the marginal distributions of the elements of $\mathbf{X}_t$ fulfill the conditions in **G2**.

The optimal projection

Ellipticity allows to compute the optimal angle ϕ. The interquantile range used in the standardization (1) can be written as $k(\tau,\alpha)\sigma_{X_j}$, where $k(\tau,\alpha)$ is a non-random positive constant. Likewise $\mathrm{IQR}_k^\tau = k(\tau,\alpha)\sigma_{X_k}$. Then, by **E1** the bivariate random vector $Y^{j\,k} = (Y_{j\,t}, Y_{k\,t})$ is elliptically distributed with a 2×2 shape matrix $k(\tau,\alpha)^{-1}\mathbf{R}$. The matrix $\mathbf{R}$ has diagonal elements $\rho_{11} = \rho_{22} = 1$ and off diagonal element $\rho_{12} = \sigma_{j\,k}/\sqrt{\sigma_{j\,j}\sigma_{k\,k}}$.

The probability contours of $Y^{j\,k}$ are therefore ellipsoids with axes that are in the direction of the eigenvectors of $k(\tau,\alpha)\mathbf{R}^{-1}$, and their lengths are proportional to the reciprocals of the square roots of the eigenvalues of $k(\tau,\alpha)\mathbf{R}^{-1}$. More precisely, since $\rho_{11} = \rho_{22} = 1$, the first eigenvalue is $k(\tau,\alpha)^{-1}(1+\rho_{12})$ with associated eigenvector $(1/\sqrt{2}, 1/\sqrt{2})$, while the second eigenvalue is $k(\tau,\alpha)^{-1}(1-\rho_{12})$ with associated eigenvector $(1/\sqrt{2}, -1/\sqrt{2})$.

If the relation between $X_{j\,t}$ and $X_{k\,t}$ is positive, then ρ_{12} is positive and $k(\tau,\alpha)^{-1}(1+\rho_{12})$

is the largest eigenvalue. Hence the projection in the 45-degree line

$$Z_t^{(j\,k)} = \frac{1}{\sqrt{2}}(Y_{j\,t} + Y_{k\,t}) \tag{4}$$

is optimal. If the relation is negative, the projection in the 135-degree line

$$Z_t^{(j\,k)} = \frac{1}{\sqrt{2}}(Y_{j\,t} - Y_{k\,t}) \tag{5}$$

is optimal. In either way, $Z_t^{(j\,k)}$ is the first principal component of $Y^{j\,k}$.

The choice between the 45- and the 135-degree lines depends on the sign of ρ_{12}. If wrongly chosen, conclusions may be misleading, as exemplified in Figure 4. Panel (a) of the figure shows the projection on the 45-degree line when the relation is negative. The projection $Z^{(j\,k)}$ is concentrated around the origin and $\text{TailCoR}^{(j\,k)\,\xi}$ is smaller than one, leading to the false conclusion that tails are thinner than Gaussian. Projecting in the 135-degree line, as shown in panel (b), captures correctly the negative relation between $Y_{j\,t}$ and $Y_{k\,t}$.

[FIGURE 4 ABOUT HERE]

In practice, we suggest the following simple algorithm for choosing between the 45- and the 135-degree lines, which is based on the fact that under independence the IQRs of both projections are the same. Let $Z_{45\,t}^{(j\,k)}$ and $Z_{135\,t}^{(j\,k)}$ be the corresponding projections. If $\text{IQR}_{45}^{(j\,k)\,\xi} > \text{IQR}_{135}^{(j\,k)\,\xi}$ then $Z_{45\,t}^{(j\,k)}$ is used, otherwise $Z_{135\,t}^{(j\,k)}$ is selected.

TailCoR under ellipticity

We are now ready to disentangle $\text{TailCoR}^{(j\,k)\,\xi}$ into the linear and non-linear components.

Theorem 1 Let $X_{j\,t}$ and $X_{k\,t}$ be two elements of the random vector $\mathbf{X}_t$ that fulfills assumptions **E1**-**E5**. Let $\rho_{j\,k}$ be the linear correlation, and let $s(\xi, \tau, \alpha)$ be a continuous function of α. Then

$$\text{TailCoR}^{(j\,k)\,\xi} = s_g(\xi, \tau) s(\xi, \tau, \alpha) \sqrt{1 + |\rho_{j\,k}|},$$

where $s_g(\xi,\tau)$ is a normalization such that under independence $\text{TailCoR}^{(j\,k)\,\xi} = 1$, the reference value.

Proof See Appendix P.

The rightmost element, $\sqrt{1+|\rho_{j\,k}|}$, captures the linear contribution to $\text{TailCoR}^{(j\,k)\,\xi}$, while $s(\xi,\tau,\alpha)$ captures the non-linear contribution as it depends on the tail index α. We will denote these contributions as linear and non-linear correlations.

Panel (a) of Figure 2 displays $s_g(\xi,\tau)s(\xi,\tau,\alpha)$ as a function of α and assuming a Student-t distribution (and for $\xi = 0.95$ and $\tau = 0.75$). The tail index varies from 2.5 to 30. The non-linear correlation decreases as α increases, and it approaches to 1 as α goes to 30, or when the distribution is indistinguishable from the Gaussian. Panel (b) shows the sensitivity of $s_g(\xi,\tau)s(\xi,\tau,\alpha)$ to ρ (with $\alpha = 2.5$). The non-linear correlation is not affected by ρ, which confirms that the two components capture different aspects of $\text{TailCoR}^{(j\,k)\,\xi}$.

[FIGURE 2 ABOUT HERE]

$\text{TailCoR}^{(j\,k)\,\xi}$ has numerous properties. First, it captures non-linear correlations if tails deviate from Gaussianity, i.e. $s(\xi,\tau,\alpha)s_g(\xi,\tau)$ can be greater, equal, or smaller than 1. Second, even if X_j and X_k are linearly uncorrelated, $\text{TailCoR}^{(j\,k)\,\xi}$ is different from one if X_j and X_k are non-Gaussian. This is akin to the coefficient of tail dependence steaming from copula theory.[8] Third, if $\mathbf{X}_t$ is Gaussian, $\text{TailCoR}^{(j\,k)\,\xi} = \sqrt{1+|\rho_{j\,k}|}$, i.e. the only source of correlation is linear.

This last property shows that under Gaussianity the upper bound is $\sqrt{2}$. Otherwise the upper bound depends on α and ξ. Figure 3 displays similar curves to those in panel (a) of Figure 2 for values of ξ typically used in practice: 0.90 (solid line), 0.95 (thick dashes), 0.975 (thin dashes), and 0.99 (tick and thin dashes). The further we explore the tails, the larger is $s_g(\xi,\tau)s(\xi,\tau,\alpha)$, and so is TailCoR. Note that the non-linear correlation induced by the heaviness of the tails translates into values of $s_g(\xi,\tau)s(\xi,\tau,\alpha)$ that are in the range $1-1.3$

[8] For instance, the tail dependence coefficient of a bivariate Student-t copula is $2t_{\alpha+1}\left(-\sqrt{\frac{(\alpha+1)(1-\rho)}{1+\rho}}\right)$, where $t_{\alpha+1}(\cdot)$ is a standardized Student-t cumulative distribution function with tail index $\alpha+1$. Even if $\rho = 0$ tail dependence is positive (unless $\alpha \to \infty$).

(except if ξ is close to one and tails are very heavy), which translates into a range of values of TailCoR between 1 and 1.83 (the latter is for $\rho = 1$ and $s_g(\xi,\tau)s(\xi,\tau,\alpha)$=1.3).

We close this section with the issue that TailCoR is not a correlation in the traditional sense, as it is not centered at zero nor bounded above. The following alternative expression:

$$\text{TailCoR}_{alt}^{(j\,k)\,\xi} = \text{sign}(\rho_{j\,k})\frac{\text{TailCoR}^{(j\,k)\,\xi} - 1}{s_g(\xi,\tau)s(\xi,\tau,\alpha)\sqrt{2} - 1} \tag{6}$$

is bounded above and below by 1 and -1, and centered around zero. Indeed, if $\rho_{j\,k} = 0$, (6) becomes

$$\text{TailCoR}_{alt}^{(j\,k)\,\xi} = \text{sign}(\rho_{j\,k})\frac{s_g(\xi,\tau)s(\xi,\tau,\alpha) - 1}{s_g(\xi,\tau)s(\xi,\tau,\alpha)\sqrt{2} - 1},$$

which is 0 only under Gaussianity; otherwise it is between -1 and 1. If $X_{j\,t}$ and $X_{k\,t}$ are perfectly correlated, $|\rho_{j\,k}| = 1$ and (6) equals 1 if $\rho_{j\,k} = 1$, and -1 if $\rho_{j\,k} = -1$.

2.3 Estimation

Estimation under **G1-G2** is straightforward and is divided in three simple steps.

Step 1 Standardize $X_{j\,t}$ and $X_{k\,t}$ and compute the projection $Z_t^{(j\,k)}$.

Step 2 Estimate the IQR of the projection: $\hat{\text{IQR}}_T^{(j\,k)\,\xi}$.

Step 3 Choose the normalization $s_g(\xi,\tau)$ and compute $\hat{\text{TailCoR}}_T^{(j\,k)\,\xi} = s_g(\xi,\tau)\hat{\text{IQR}}_T^{(j\,k)\,\xi}$.

Under **E1-E5** the linear correlation $\rho_{j\,k}$ is estimated with a robust method. Lindskog et al. (2003) introduce a robust estimator that exploits the geometry of the elliptical distribution. Let $\hat{\kappa}_{j\,k,T}$ be the estimator of the Kendall's correlation. Lindskog et al. (2003) show that $\hat{\kappa}_{j\,k,T}$ is invariant in the class of elliptical distributions. Then

$$\hat{\rho}_{j\,k,T} = \sin\left(\frac{\pi}{2}\hat{\kappa}_{j\,k,T}\right),$$

and $\sqrt{1+\hat{\rho}_{j\,k,T}}$ follows. Given the $\hat{\text{IQR}}_T^{(j\,k)\,\xi}$ obtained in step 2 above, the estimator of the non-linear correlation is

$$\hat{s}(\xi,\tau,\alpha)_T = \frac{\hat{\text{IQR}}_T^{(j\,k)\,\xi}}{\sqrt{1+\hat{\rho}_{j\,k,T}}}.$$

We now see the computational advantages of $\text{TailCoR}^{(j\,k)\,\xi}$. It can be estimated exactly for any probability level ξ, and no optimizations are needed as it is based on simple steps, each requiring a few lines of programming code. This makes $\text{TailCoR}^{(j\,k)\,\xi}$ fast to compute. Moreover, the estimation of the tail index is not required.

The following theorem shows the asymptotic properties of $\hat{\text{TailCoR}}_T^{(j\,k)\,\xi}$.

Theorem 2 Let $X_{j\,t}$ and $X_{k\,t}$ be two elements of the random vector $\mathbf{X}_t$ that fulfills assumptions **E1** – **E5**. Let $f_{(j\,k)}(\cdot)$ and $F_{(j\,k)}(\cdot)$ be the probability and cumulative density functions of $Z_t^{(j\,k)}$, and let $I_{\{.\}}$ be an indicator function that takes value 1 if its argument is true. Then, as $T \to \infty$

$$\sqrt{T}\left(\hat{\text{TailCoR}}_T^{(j\,k)\,\xi} - \text{TailCoR}^{(j\,k)\,\xi}\right) \to_d \mathcal{N}\left(0, 4s_g(\xi,\tau)^2 \frac{\Gamma(Q^{(j\,k)\xi})}{f^2_{(j\,k)}(F^{-1}_{(j\,k)}(\xi))}\right),$$

where

$$\Gamma(Q^{(j\,k)\xi}) = \sum_{t=-\infty}^{+\infty} E\left(W_0(Q^{(j\,k)\xi})W_t(Q^{(j\,k)\xi})\right) \text{ and}$$

$$W_t(Q^{(j\,k)\xi}) = I_{\{Z_t^{(j\,k)} \leq Q^{(j\,k)\xi}\}} - P(Z_t^{(j\,k)} \leq Q^{(j\,k)\xi}).$$

Proof See Appendix P.

Five remarks to the theorem. First, the univariate density $f_{(j\,k)}(\cdot)$ in the denominator is elliptical, and therefore easy to compute. Alternatively, the asymptotic variance can be computed by bootstrap. Second, $\Gamma(Q^{(j\,k)\xi})$ is the long-run component that accounts for the time dependence. Third, it is possible to skip **E3** and derive the elliptical representation of TailCoR without moments. Then $\rho_{j\,k}$ becomes the $(j\,k)$ element of the standardized dispersion matrix that can be estimated with the Tyler's M-estimator (Tyler (1987)). Fourth, admittedly the theorem ignores the effect of the estimated median and IQR in the standardization of $X_{j\,t}$ and $X_{k\,t}$ (step 1). Monte Carlo results –available upon request– indicate that their effect is negligible. Last, it is possible to derive an equivalent asymptotic distribution under **G1**-**G2**, as shown in the following corollary.

Corollary 1 Let $X_{j\,t}$ and $X_{k\,t}$ be two elements of the random vector $\mathbf{X}_t$ that fulfills assumptions **G1** – **G2**. Let $f_{(j\,k)}(\cdot)$ and $F_{(j\,k)}(\cdot)$ be the probability and cumulative density functions of $Z_t^{(j\,k)}$, and let $I_{\{\cdot\}}$ be an indicator function that takes value 1 if its argument is true. Then, as $T \to \infty$

$$\sqrt{T}\left(\hat{\text{TailCoR}}_T^{(j\,k)\,\xi} - \text{TailCoR}^{(j\,k)\,\xi}\right) \to_d \mathcal{N}\left(0, s_g(\xi,\tau)^2 \Upsilon\right),$$

where

$$\Upsilon = \frac{\Gamma(Q^{(j\,k)\xi})}{f^2_{(j\,k)}(F^{-1}_{(j\,k)}(\xi))} + \frac{\Gamma(Q^{(j\,k)1-\xi})}{f^2_{(j\,k)}(F^{-1}_{(j\,k)}(1-\xi))} - 2\frac{\Gamma(Q^{(j\,k)\xi}, Q^{(j\,k)1-\xi})}{f_{(j\,k)}(F^{-1}_{(j\,k)}(\xi)) f_{(j\,k)}(F^{-1}_{(j\,k)}(1-\xi))},$$

$$\Gamma(Q^{(j\,k)\xi}, Q^{(j\,k)1-\xi}) = \sum_{t=-\infty}^{+\infty} E\left(W_0(Q^{(j\,k)\xi}) W_t(Q^{(j\,k)1-\xi})\right), \text{ and}$$

$$W_t(Q^{(j\,k)\xi}) = I_{\{Z_t^{(j\,k)} \leq Q^{(j\,k)\xi}\}} - P(Z_t^{(j\,k)} \leq Q^{(j\,k)\xi}).$$

Proof If follows the proof of Theorem 2 but $Q^{(j\,k)\xi} \neq Q^{(j\,k)1-\xi}$.

2.4 Multivariate TailCoR

So far we only considered the pair $(j\,k)$ of random variables while the random vector $\mathbf{X}_t$ is of dimension N. Considering all the pairs leads to a $N(N+1)/2 \times N(N+1)/2$ symmetric matrix of TailCoRs (including TailCoR of a random variable with itself). For the ease of exposition let $\tilde{N} = N(N+1)/2$. We denote by $\xi_{(j\,k)}$ the probability level at which we compute the IQR for the $(j\,k)$ projection and we assume $\xi_{(j\,k)} = \xi \ \forall j,k$. We define the matrix of TailCoR as

Definition 2 Under **G1-G2**, the matrix of TailCoRs is defined as follows

$$\mathbf{TailCoR}^{\xi} := s_g(\xi,\tau)\mathbf{IQR}^{\xi}, \tag{7}$$

where $\mathbf{IQR}^{\xi}$ is a matrix of IQRs of the $\tilde{N} \times 1$ projections.

The assumption $\xi_{(j\,k)} = \xi\ \forall j,k$ is a simplification that allows definition (7), but it is possible to relax it at the expense of notation. Under **E1**-**E5**, (7) becomes

$$\mathbf{TailCoR}^{\xi} = \sqrt{2} s_g(\xi,\tau) s(\xi,\tau,\alpha) \mathbf{\Psi}, \tag{8}$$

where the matrix $\mathbf{\Psi}$ has $(j\,k)$ element $\sqrt{\frac{1+|\rho_{j\,k}|}{2}}$. This matrix is symmetric, with unitary diagonal elements, and off-diagonal elements bounded above and bellow by 1 and $\sqrt{1/2}$ respectively. This matrix is invariant to location-scale shifts of $\mathbf{X}_t$, and positive definite.

Estimation follows the same steps as in the univariate case under **G1**-**G2**: $\hat{\mathbf{TailCoR}}^{\xi}_T = s_g(\xi,\tau)\hat{\mathbf{IQR}}^{\xi}$. Under **E1**-**E5**, an extra step has to be added as $s(\xi,\tau,\alpha)$ is the same for all pairs. Let $\hat{s}(\xi,\tau,\alpha)_{h,T} = \hat{s}(\xi,\tau,\alpha)^{(j\,k)\,T}$, $h = 1,\ldots,\tilde{N}$. The non-linear correlation is estimated by pooling the pairwise estimators:

$$\hat{s}(\xi,\tau,\alpha)_T = \frac{1}{\tilde{N}} \sum_{h=1}^{\tilde{N}} \hat{s}(\xi,\tau,\alpha)_{h,T}.$$

Estimating tail indices by cross-sectionally averaging can be found in, for instance, Nolan (2013) and Dominicy et al. (2013).

The asymptotic distribution requires the use of results in Dominicy et al. (2013) on covariances between the sample quantiles of marginal distributions under S-mixing. The following theorem shows the asymptotic distribution of the vectorized $\hat{\mathbf{TailCoR}}^{\xi}_T$.

Theorem 3 Let $\mathbf{X}_t$ be a random vector that fulfills assumptions **E1** – **E5**. Let $vech\hat{\mathbf{TailCoR}}^{\xi}_T$ be the $\frac{\tilde{N}(\tilde{N}+1)}{2} \times 1$ half-vectorized matrix of $\hat{\mathbf{TailCoR}}^{\xi}_T$. Likewise for $vech\mathbf{TailCoR}^{\xi}$. Let $f_j(\cdot)$ and $F_j(\cdot)$ be the probability and cumulative density functions of the jth projection ($j = 1,\ldots,\frac{\tilde{N}(\tilde{N}+1)}{2}$), and let $I_{\{\cdot\}}$ be an indicator function that takes value 1 if its argument is true. Then, as $T \to \infty$

$$\sqrt{T}\left(vech\hat{\mathbf{TailCoR}}^{\xi}_T - vech\mathbf{TailCoR}^{\xi}\right) \to \mathcal{N}\left(0, 4 s_g(\xi,\tau)^2 \mathbf{\Omega}\right),$$

where $\mathbf{\Omega}$ is a $\frac{\tilde{N}(\tilde{N}+1)}{2} \times \frac{\tilde{N}(\tilde{N}+1)}{2}$ matrix with $(j\,j)$ element

$$\begin{aligned}
\Omega_{j\,j} &= \frac{\Gamma_{j\,j}(Q_j^{\xi})}{f_j^2(F_j^{-1}(\xi))}, \text{ where} \\
\Gamma_{j\,j}(Q_j^{\xi}) &= \sum_{t=-\infty}^{\infty} E(W_0(Q_j^{\xi}), W_t(Q_j^{\xi})) \text{ and} \\
W_t(Q_j^{\xi}) &= I_{\{Z_{j\,t} \leq Q_j^{\xi}\}} - P(Z_{j\,t} \leq Q_j^{\xi}).
\end{aligned}$$

The $(j\,k)$ element of $\mathbf{\Omega}$ is

$$\begin{aligned}
\Omega_{j\,k} &= \frac{\Gamma_{j\,k}(Q_j^{\xi}, Q_k^{\xi})}{f_j(F_j^{-1}(\xi)) f_k(F_k^{-1}(\xi))} \forall j \neq k, \text{ where} \\
\Gamma_{j\,k}(Q_j^{\xi}) &= \sum_{t=-\infty}^{\infty} E(W_0(Q_j^{\xi}, Q_k^{\xi}), W_t(Q_j^{\xi}, Q_k^{\xi})) \text{ and} \\
W_t(Q_j^{\xi}, Q_k^{\xi}) &= I_{\{Z_{j\,t} \leq Q_j^{\xi}, Z_{k\,t} \leq Q_k^{\xi}\}} - P(Z_{j\,t} \leq Q^{\xi}, Z_{k\,t} \leq Q_k^{\xi}).
\end{aligned}$$

Proof See Appendix P.

3 A Monte Carlo study

We analyze the finite sample properties of TailCoR with a Monte Carlo study. We consider three bivariate elliptical distributions: Gaussian, Student-t with $\alpha = 2.5$, and ES with $\alpha = 1.5$. The most heavy tailed distribution is the Student-t, followed by the ES and the Gaussian. The location parameters are set to zero and the dispersion matrix has unitary diagonal elements and off-diagonal elements 0.50. We consider three samples sizes $T = \{100, 1000, 5000\}$ and two sets of replications $H = \{100, 500\}$. In the sequel we show results for $T = 5000$ and $H = 500$, except for the last table and figure. Results for other configurations are alike and available under request.

[FIGURE 5 ABOUT HERE]

Figure 5 shows the finite sample distributions of the TailCoR estimates for $\xi = 0.95$ and for the three distributions (solid line for the Gaussian, dashed for the Student-t, and dotted for

the ES). In all cases, TailCoR is larger than one. The estimated TailCoR is more precise under Gaussianity than under heavy tails, as it only depends on the linear correlation. Moreover, the median is around 1.22, very close to the true value of the linear correlation $\sqrt{1+0.50} = 1.225$. By contrast, estimators under the Student-t and the ES have medians well above, 1.42 for the ES and 1.47 for the Student-t, reflecting the non-linear dependencies.

[FIGURE 6 ABOUT HERE]

Figure 6 shows the sensitivity of TailCoR to ξ for the Student-t with $\alpha = 2.5$ (panel a), the ES with $\alpha = 1.5$ (panel b), and the Gaussian (panel c). Each line is the density of the 500 estimates of TailCoR for different values of ξ: 0.90 (solid line), 0.95 (dashed) and 0.99 (dotted). The densities overlap for the Gaussian distribution since $\text{TailCoR}^{(j\,k)\,\xi} = \sqrt{1+\rho_{j\,k}}$ does not depend on ξ; the small differences are due to finite sample discrepancies. Regarding the other distributions, results show that TailCoR increases with ξ as we explore further the tails.

[FIGURE 7 ABOUT HERE]

The precision and convergence in distribution of the estimator are shown in Table 2 and Figure 7 for $\xi = 0.95$. For each combination of T and H the table shows the average and variance of the H estimated TailCoRs. The mean of the estimates is quite stable across the sample size and the number of replications, while the precision increases with both T and H, though the sample size is clearly more important. The solid line in the Figure is the standardized Gaussian while the bars are the histograms of the standardized estimates for different sample sizes and replications (indicated at the top of each plot). As expected, the histograms approach to the limiting distribution as the sample size and the replications increase.

4 Downside and upside TailCoR under asymmetry

It is often the case that the interest lies in a particular region of the distribution. For instance, in risk management the region of interest is where $Y_{j\,t}$ and $Y_{k\,t}$ are negative. This translates

into the negative side of the distribution of the projection $Z_t^{(j\,k)}$, which leads to the following definition.

Definition 3 Under **G1-G2** we define downside- and upside-TailCoR as follows:

$$\begin{aligned} \text{TailCoR}^{(j\,k)\,\xi-} &:= 2s_g(\xi,\tau)\text{IQR}^{(j\,k)\,\xi-} \text{ and} \\ \text{TailCoR}^{(j\,k)\,\xi+} &:= 2s_g(\xi,\tau)\text{IQR}^{(j\,k)\,\xi+}. \end{aligned}$$

$\text{IQR}^{(j\,k)\,\xi-} = \text{Q}^{(j\,k)\,0.50} - \text{Q}^{(j\,k)\,1-\xi}$ is the difference between the median and the lower tail quantile. Similarly $\text{IQR}^{(j\,k)\,\xi+} = \text{Q}^{(j\,k)\,\xi} - \text{Q}^{(j\,k)\,0.50}$ is the difference between the upper tail quantile and the median.

Note that the scale 2 in the definition makes $\text{TailCoR}^{(j\,k)\,\xi-}$ and $\text{TailCoR}^{(j\,k)\,\xi+}$ comparable with the full fledge TailCoR. Note also that under ellipticity, the projection $Z_t^{(j\,k)}$ is symmetric and downside and upside TailCoRs are equal to the full fledge TailCoR. The value added of downside and upside TailCoRs is therefore put forward under asymmetry.

We model asymmetry with a normal mean-variance mixture (or NMVM, see e.g. Mencia and Sentana (2009) and Jondeau (2010)), i.e. $\mathbf{X}_t =_d \boldsymbol{\mu} + \mathcal{R}_{\alpha\,t}^2\boldsymbol{\gamma} + \mathcal{R}_{\alpha\,t}\boldsymbol{\Lambda}\boldsymbol{\Upsilon}_t$, where $\boldsymbol{\Upsilon}_t$ is a standardized multivariate Gaussian distribution, and $\boldsymbol{\gamma}$ is a vector of parameters that capture the asymmetry. The most prominent distribution that belongs to this family is the generalized hyperbolic, as it nests the hyperbolic, normal gamma, and normal inverse Gaussian, among others. As the elliptical family, the NMVM family enjoys the property of affine invariance, i.e. it is closed under linear operations, which means that the projection $Z_t^{(j\,k)}$ is also a NMVM random variable.

The vector of unknown parameters is now $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}, \alpha)$. While assumption **E1** has to be trivially modified, **E2-E5** remain unchanged. We re-name them as **A1-A5**. In order to disentangle the linear and non-linear correlations, uniform asymmetry is needed, as stated in the following assumption.

A6 $\gamma = \gamma_j = \gamma_k$, $\forall j, k = 1, \ldots, N$.

This assumption does not imply symmetry, which is a particular case when $\gamma = 0$. Rather it implies that all the random variables are equally asymmetric. Though uniform asymmetry may be seen restrictive, empirical results – see for instance Jondeau (2010) – show that the estimated asymmetries for the random variables are around zero without large discrepancies between them.

Under asymmetry the relation between $X_{j\,t}$ and $X_{k\,t}$ is not the same over the domain of the observations. For instance, Longin and Solnik (2001) and Jondeau (2010) find that the dependence across asset returns may be stronger in bearish than in bullish conditions. These asymmetric dependencies are captured with semi variances, semi interquantile ranges, and semi correlations. The term semi relates to the use of observations that are in a certain range. For instance, the positive semi interquantile range is the difference between the upper quantile and the median, as used in the upside-TailCoR.

Let $\sigma_j^{2\,+}$ be the positive semi variance and $\mathrm{IQR}_j^{\tau\,+} = \mathrm{Q}_j^{\tau} - \mathrm{Q}_j^{0.50}$ be the positive semi interquantile range of $X_{j\,t}$. The relation between both is given by $\mathrm{IQR}_j^{\tau\,+} = k(\tau, \alpha, \gamma)^{+}\sigma_j^{+}$. Similarly, $\sigma_j^{2\,-}$ and $\mathrm{IQR}_j^{\tau\,-} = \mathrm{Q}_j^{0.50} - \mathrm{Q}_j^{1-\tau}$ are the negative semi variance and the negative semi interquantile range, which are related by $\mathrm{IQR}_j^{\tau\,-} = k(\tau, \alpha, \gamma)^{-}\sigma_j^{-}$. The same definitions apply for $X_{k\,t}$. We also need the positive and negative semi correlations between $X_{j\,t}$ and $X_{k\,t}$. The former is defined as $\rho_{j\,k}^{+} = \sigma_{j\,k}^{+}/\sigma_j^{+}\sigma_k^{+}$, where $\sigma_{j\,k}^{+}$ is the positive semi covariance. Likewise for the negative semi correlation.

By the affine invariance of the NMVM family, it follows that semi interquantile ranges of previous definition equal $\mathrm{IQR}^{(j\,k)\,\xi\,+} = k(\xi, \alpha, \gamma)^{+}\sigma_{(j\,k)}^{+}$ and $\mathrm{IQR}^{(j\,k)\,\xi\,-} = k(\xi, \alpha, \gamma)^{-}\sigma_{(j\,k)}^{-}$, where $\sigma_{(j\,k)}^{2\,+}$ and $\sigma_{(j\,k)}^{2\,-}$ are the positive and negative semi variances of the projection $Z_t^{(j\,k)}$. Further properties of the semi moments are found in Appendix P.

Under the NMVM family of distributions the optimal projection may not be on the 45- or the 135-degree lines, but on a line with angle ϕ and that maximizes the variability of the projection. The following theorem shows downside- and upside-TailCoR under asymmetry.

Theorem 4 Let $X_{j\,t}$ and $X_{k\,t}$ be two elements of the random vector $\mathbf{X}_t$ that fulfills assumptions **A1**-**A6**. Let $\rho_{j\,k}^{+}$ and $\rho_{j\,k}^{-}$ be the positive and negative semi correlations, and let

$s(\xi,\tau,\alpha,\gamma)^+$ and $s(\xi,\tau,\alpha,\gamma)^-$ be two continuous functions of α. Then

$$\begin{aligned}\text{TailCoR}^{(j\,k)\,\xi\,-} &= 2s_g(\xi,\tau)s(\xi,\tau,\alpha,\gamma)^-\sqrt{1+2|\rho^-_{j\,k}|\sin\phi\cos\phi}\\ \text{TailCoR}^{(j\,k)\,\xi\,+} &= 2s_g(\xi,\tau)s(\xi,\tau,\alpha,\gamma)^+\sqrt{1+2|\rho^+_{j\,k}|\sin\phi\cos\phi} \quad \text{and,}\end{aligned}$$

where $s_g(\xi,\tau)$ is a normalization such that under independence $\text{TailCoR}^{(j\,k)\,\xi}=1$, the reference value.

Proof See Appendix P.

Note that when ϕ equals $\frac{\pi}{4}$ or $\frac{3\pi}{4}$ (i.e. projections in the 45- and 135-degree lines), $\sqrt{1+2\sin\phi\cos\phi|\rho^+_{j\,k}|}$ equals $\sqrt{1+|\rho^+_{j\,k}|}$. The downside- and upside-TailCoR can differ because either the linear and/or the non-linear correlations are different. Estimation follows the same steps as in the previous estimators except for the use of semi metrics.

5 TailCoR of eight major US banks

We illustrate TailCoR with an application to 10 years of daily returns of eight major US banks: Bank of America (BAC), Bank of New York (BK), Citigroup (C), Goldman Sachs (GS), JP Morgan (JPM), Morgan Stanley (MS), Wells Fargo (WFC), and State Street (STT). The sample starts in January 2, 2003 and ends in December 31, 2012.

[TABLE 3 ABOUT HERE]

Table 3 shows descriptive statistics. All the metrics are quantile-based. The column kurtosis is a measure of excess kurtosis, computed as $\text{IQR}^{0.975}/\text{IQR}^{0.75}-2.91$. The skewness is computed as $(Q^{0.975}-Q^{0.50})-(Q^{0.50}-Q^{0.025})$. Visual inspection –not shown here– reveals strong volatility clustering, which is a conditional feature that may distort the unconditional measurement of the tails and their correlations. To safeguard against it, we adjust the returns with a GARCH(1,1). The table shows that the adjusted returns share the stylized facts of financial returns: zero medians, a great deal of fat tails, and they are approximately

symmetric.[9]

Next, we show full-sample results and, to study the evolution through time, estimations using biannual and non-overlapping windows. For comparison we also compute upper and lower exceedance correlations (denoted $\theta^+(>\xi)$ and $\theta^-(<\xi)$ respectively), and the parametric and non-parametric tail dependence coefficients (the former with a copula-t and the latter as in Straetmans et al. (2008)). We denote them τ_p and the τ_{np} respectively. Results for TailCoR and positive exceedance correlation are for $\xi = 0.975$ ($\xi = 0.025$ for downside correlation). Results for τ_{np} are for $\xi = 0.025$. Estimations for other values of ξ are available upon request.[10]

[TABLE 4 ABOUT HERE]

Upper panel of Table 4 shows the matrix of TailCoRs and their standard deviations in parenthesis. They are all above one, the reference value, indicating that, as expected, returns are far from being Gaussian and uncorrelated. Also, they have very low standard errors and a simple t-stat would conclude that they are all significantly larger than one. The diagonal elements are the TailCoRs of a bank with itself. They are among the largest of the matrix, as they can be interpreted as tail risks. The maximum off-diagonal element is 1.64 and the minimum is 1.43, a tight interval that reflects the strong interrelation between the banks.

In order to understand the contributions of the linear and non-linear correlations, middle and bottom panels of the table show the matrices of $\sqrt{1+\rho}$ and $s_g(0.975, 0.75)s(0.975, 0.75, \alpha)$ respectively (and their standard errors in parenthesis). The diagonal elements of the middle panel are equal to 1.41, or $\sqrt{2}$. The off-diagonal elements are between 1.32 and 1.39 (or Pearson correlations between 0.76 and 0.94). Many standard deviations are 0.00. They are not numerically 0 (except in the main diagonal), but smaller than 0.00, reflecting that estimates are very precise. The bottom panel is the most interesting. All the non-linear contributions are above one and precisely estimated, and the diagonal elements are among the largest of the matrix. If, hypothetically, the adjusted returns would be Student-t distributed, $s_g(0.975, 0.75)s(0.975, 0.75, \alpha) = 1.13$ corresponds to a tail index of approximately seven,

[9] We also did the analysis without the GARCH adjustment. Results are very similar –because of the standardization (1)– and are available upon request.

[10] Since τ_{np} depends on the Hill statistic, alternative methods for choosing ξ are available. See Straetmans et al. (2008) for a discussion.

evincing the intrinsic heaviness of the tails and the significant probability of joint extreme values. As it will be verified later, these high values are due to a large number of extreme observations (both in frequency and magnitude) at the height of the financial crisis.

[TABLE 5 ABOUT HERE]

Equivalent matrices for the upside and downside exceedance correlations, and the tail dependence coefficients are shown in Appendix R. Results read along the same lines as TailCoR. Table 5 displays the matrix of correlations between TailCoR$^{0.975}$, $\theta^{+}(> 0.975)$, $\theta^{-}(< 0.025)$, τ_p, and τ_{np}, as well as with the linear and non-linear contributions. The correlations are computed by half-vectorizing the fives matrices, and calculating the Pearson correlations between them. The correlation between $s(0.975, 0.75, \alpha)$ and $\sqrt{1+\rho}$ is 0.07, as expected and confirming that the linear and non-linear correlations capture different features of the dependence between returns. Moreover, TailCoR seems to be driven more by the non-linear than by the linear contribution (correlations of 0.89 and 0.52 respectively), which is also found in the dynamic estimations below. The correlations of TailCoR with the other measures of dependence are high, in the range 0.72-0.75, reflecting that all the measures capture the same features of the panel of returns. While this is reassuring, biannual estimations, shown below, illlustrate the advantage of TailCoR over exceedance correlation and coefficients of tail dependence in small and medium samples.

[FIGURE 8 ABOUT HERE]

Indeed, panels (a), (b) and (c) of Figure 8 display the dynamic evolution of TailCoR, τ_p, and $\tau_n p$ respectively and for all pair of banks. The exceedance correlations could not be computed due to lack of observations. The coefficients τ_p and τ_{np} are quite erratic and differ substantially. The former often hits zero, which implies Gaussianity and is at odds with the high degree of co-movement and heavy tails described in Table 4. The latter wiggles in a tight interval around 98.2%. By contrast, TailCoR is more stable and shows a pattern more in line with what it is expected from the financial and economic events that happened during the sample period.

Figure 9 focuses on TailCoR (panel a) and its linear (panel b) and non-linear contributions (panel c). For the shake of visibility and interpretation, each line is the cross-sectional average of one bank with respect to the others. For instance, the value of TailCoR for BAC in the first half of 2003 is the average TailCoR of BAC with respect to all the other banks on that period.

TailCoR is heterogeneous during 2003-2004, with estimates that range from heavy tailed (above 1.2) to Gaussian-like (around 1). At the beginning of 2005, TailCoR concentrates around value near 1, in a period characterized by Gaussian tails, and in some cases even thinner (see same period in panel b). Since then, the heaviness and the co-movements of the tails increase steadily with a pattern that is clearly identified with financial, economic and political events. March 2007 has been identified as the beginning of the financial crisis. Acharya and Richardson (2009) identify March 5 as the first date in the timeline of the crisis. Then TailCoR decreases during the second part of the year and it picks up again in 2008, when further troubles in the financial industry were made public: Citigroup and Merrill Lynch announced pronounced losses, the UK Government nationalizes Northern Rock, AIG announces troubles in the valuation of CDS, and in September 15 Lehman Brothers files for bankruptcy. TailCoR then lowers smoothly until 2010. Early that year there were increasing concerns over emerging problems in European sovereign debt markets. Though the first half of 2012 is a relatively calm period for European debt markets, the last 6 months of the sample show a sudden and important increase in TailCoR, to levels similar to those of the height of the financial crisis.

[FIGURE 9 ABOUT HERE]

The non-linear contributions (panel b) show a similar pattern to TailCoR while the linear contributions (panel c) are smoother across time, which implies that the variations of TailCoR are mainly driven by the tails. The last part of the sample is revealing: while the linear contributions show a decline, the non-linear exhibit a sharp increase. This upward movement shows that the risks associated with the potential exit of Greece from the euro area, the political uncertainty in Italy, and the troubles in the Spanish financial sector that lead to the

bailout, weighed in the tail correlations of the major US banks.

6 Conclusions

We have introduced TailCoR, a new metric for tail correlations that is a function of linear and non-linear contributions, the latter characterized by the heaviness of the tails. TailCoR is exact for any probability level, it does not depend of any specific distributional assumption, and no optimizations are needed. Monte Carlo simulations reveal its goodness in finite samples. An empirical illustration to a panel of eight major US banks shows an increase of TailCoR during the financial crisis that was due to a surge of linear and non-linear correlations. The end of 2012 also shows an increase of TailCoR, which is only driven by the non-linear correlation, reflecting the tail risks steaming from Europe.

Several extensions are possible. One is dynamic TailCoR, which is based on the dynamic IQR of the projection ($\mathrm{IQR}_t^{(j\,k)\,\xi} = \mathrm{Q}_t^{(j\,k)\,1-\xi} - \mathrm{Q}_t^{(j\,k)\,\xi}$), where the time varying quantiles are regressed on the determinants of tail correlations. Under ellipticity the dynamic linear correlation $\rho_{i\,j,t}$ can be estimated with a robustified version of the DCC model (Boudt et al. (2013)). The dynamic non-linear correlation can be computed similarly to the static case: $s(\xi,\tau,\alpha)_t = \mathrm{IQR}_t^{(j\,k)\,\xi}/\sqrt{1+\rho_{j\,k,t}}$.

Acknowledgements

We are grateful to the seminar participants at ECORE, the Federal Reserve Bank of New York, the University of Valencia, the Université de Franche-Comté, Riksbank, Complutense University, HEC Lausanne, Helsinki, and the Swedish House of Finance, as well as the conference participants at the CFE11 (London, December 2011), the Humboldt-Copenhagen Conference on Recent Developments in Financial Econometrics (Berlin, March 2013), the ENTER jamboree (Brussels, March 2013), the Econometric Society Australasian Meeting (Sydney, July 2013), the First Intl. Workshop on Financial Econometrics (Natal, October 2013), and the Workshop on Skewness, Heavy Tails, Market Crashes, and Dynamics (Cambridge, April 2014).

References

Acharya, V. and M. Richardson (2009). *Restoring Financial Stability.* New Jersey: John Willey & Sons.

Beirlant, J., Y. Goegebeur, J. Segers, and J. Teugels (2004). *Statistics of Extremes. Theory and Applications.* New Jersey: John Willey & Sons.

Berkes, I., S. Hormann, and J. Schauer (2009). Asymptotic results for the empirical process of stationary sequences. *Stochastic Processes and their Applications 119*, 1298–1324.

Boudt, K., J. Danielsson, and S. Laurent (2013). Robust forecasting of dynamic conditional correlation garch models. *International Journal of Forecasting 29*, 244–257.

Chollete, L., V. de la Peña, and C.-C. Lu (2011). International diversification: A copula approach. *Journal of Banking and Finance 35*, 403–417.

Cizeau, P., M. Potters, and J.-P. Bouchaud (2001). Correlation structure of extreme stock returns. *Quantitative Finance 1*, 217–222.

Coroneo, L. and D. Veredas (2012). A simple two-component model for the distribution of intraday returns. *European Journal of Finance 18(9)*, 775–797.

Dominicy, Y., S. Hormann, H. Ogata, and D. Veredas (2013). On sample marginal quantiles for stationary processes. *Statistics and Probability Letters 83*, 28–36.

Dominicy, Y., H. Ogata, and D. Veredas (2013). Inference for vast dimensional elliptical distributions. *Computational Statistics 28*, 1853–1880.

Embrechts, P., L. de Haan, and X. Huang (2000). Modeling multivariate extremes. In P. Embrechts (Ed.), *Extremes and Integrated Risk Management*, pp. 59–67. London: Risk books.

Hartmann, P., S. Straetmans, and C. de Vries (2004). Asset market linkages in crisis periods. *Review of Economics and Statistics 86*, 313–326.

Hartmann, P., S. Straetmans, and C. de Vries (2005). Banking system stability: A cross-atlantic perspective. In M. Carey and R. Stulz (Eds.), *Risk of Financial Institutions*, pp. 133–193. Chicago: The University of Chicago Press.

Hashorva, E. (2008). Tail asymptotic results for elliptical distributions. *Insurance: Mathematics and Economics 43*, 158–164.

Hashorva, E. (2010). On the residual dependence index of elliptical distributions. *Statistics & Probability Letters 80*, 1070–1078.

Hendricks, W. and R. Koenker (1992). Hierarchical spline models for conditional quantiles and the demand for electricity. *Journal of the American Statistical Association 87*, 58–68.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. New York: Chapman & Hall/CRC.

Jondeau, E. (2010). Asymmetry in tail dependence of equity portfolios. NCCR WP 658.

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

Lindskog, F., A. McNeil, and U. Schmock (2003). Kendall's tau for elliptical distributions. In G. Bol, T. R. Gholamreza Nakhaeizadeh, Svetlozar T. Rachev, and K.-R. Vollmer (Eds.), *Credit Risk - Measurement, Evaluation and Management*. Physica–Verlag HD.

Longin, F. and B. Solnik (2001). Extreme correlation of international equity market. *Journal of Finance 56*, 649–676.

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

Mencia, J. and E. Sentana (2009). Multivariate location–scale mixtures of normals and mean–variance–skewness portfolio allocations. *Journal of Econometrics 153*, 105–121.

Nolan, J. (2013). Multivariate elliptically contoured stable distributions: theory and estimation. *Computational Statistics*, forthcoming.

Poon, S.-H., M. Rockinger, and J. Tawn (2004). Extreme value dependence in financial markets: Diagnosis, models, and financial implications. *Review of Financial Studies 17*, 581–610.

Sherman, S. (1955). A theorem on convex sets with applications. *The Annals of Mathematical Statistics 26*, 763–767.

Straetmans, S., W. Verschoor, and C. Wolf (2008). Extreme us stock market fluctuations in the wake of 9/11. *Journal of Applied Econometrics 23*, 17–42.

Tyler, D. E. (1987). A distribution–free m–estimator of multivariate scatter. *Annals of Statistics 12*, 234–251.

Appendix T: tabulation of $s_g(\xi, \tau)$

	ξ														
τ	0.700	0.725	0.750	0.775	0.800	0.825	0.850	0.875	0.900	0.925	0.950	0.975	0.990	0.995
0.600	0.483	0.424	0.375	0.335	0.301	0.271	0.244	0.220	0.198	0.176	0.154	0.129	0.109	0.098
0.625	0.607	0.533	0.472	0.422	0.379	0.341	0.307	0.277	0.249	0.221	0.194	0.163	0.137	0.124
0.650	0.735	0.644	0.571	0.510	0.458	0.412	0.372	0.335	0.301	0.268	0.234	0.196	0.166	0.150
0.675	0.865	0.759	0.673	0.601	0.539	0.486	0.438	0.394	0.354	0.315	0.276	0.231	0.195	0.176
0.700	–	0.877	0.778	0.694	0.623	0.561	0.506	0.456	0.409	0.364	0.319	0.267	0.226	0.204
0.725	–	–	0.886	0.791	0.711	0.640	0.577	0.520	0.466	0.415	0.363	0.305	0.257	0.232
0.750	–	–	–	0.893	0.801	0.722	0.651	0.586	0.526	0.468	0.410	0.344	0.290	0.262
0.775	–	–	–	–	0.898	0.808	0.729	0.657	0.589	0.525	0.459	0.385	0.325	0.293
0.800	–	–	–	–	–	0.900	0.812	0.731	0.657	0.585	0.512	0.429	0.362	0.327
0.825	–	–	–	–	–	–	0.902	0.812	0.729	0.649	0.568	0.477	0.402	0.363
0.850	–	–	–	–	–	–	–	0.901	0.809	0.720	0.630	0.529	0.445	0.402
0.875	–	–	–	–	–	–	–	–	0.898	0.799	0.699	0.587	0.494	0.447
0.900	–	–	–	–	–	–	–	–	–	0.890	0.779	0.654	0.551	0.497

Interpolation can be used for values of ξ and τ that are not in the table. Alternatively, a function can be programmed following four steps: i) simulate from a bivariate standardized and uncorrelated Gaussian, ii) standardize the variables: $Y_{1\,t} = (X_{1\,t} - \mathrm{Q}_1^{0.50})/\mathrm{IQR}_1^{\tau}$ and $Y_{2\,t} = (X_{2\,t} - \mathrm{Q}_2^{0.50})/\mathrm{IQR}_2^{\tau}$, iii) compute the projection $Z_t^{(1\,2)}$ and $\mathrm{IQR}^{(1\,2)\,\xi}$, and iv) $s_g(\xi, \tau) = 1/\mathrm{IQR}^{(1\,2)\,\xi}$. These steps can be repeated many times and the median normalization is used.

Appendix P: proofs

Proof of Theorem 1

We first consider the case when $\rho_{j\,k} > 0$ and hence $Z_t^{(j\,k)} = \frac{1}{\sqrt{2}}(Y_{j\,t} + Y_{k\,t})$. The variance of $Y_{j\,t}$ is

$$\sigma^2_{Y_j} = \frac{\sigma^2_{X_j}}{(\mathrm{IQR}^\tau_j)^2},$$

and likewise for $Y_{k\,t}$. The variance of $Z_t^{(j\,k)}$ is

$$\sigma^2_{(j\,k)} = \frac{1}{2}\left(\frac{\sigma^2_{X_j}}{(\mathrm{IQR}^\tau_j)^2} + \frac{\sigma^2_{X_k}}{(\mathrm{IQR}^\tau_k)^2} + 2\sigma_{Y_j Y_k}\right),$$

where $\sigma_{Y_j Y_k}$ is the covariance between $Y_{j\,t}$ and $Y_{k\,t}$. Since $\mathrm{IQR}^\tau_j = k(\tau,\alpha)\sigma_{X_j}$ and $\mathrm{IQR}^\tau_k = k(\tau,\alpha)\sigma_{X_k}$

$$\sigma^2_{(j\,k)} = \frac{1}{2}\left(\frac{\sigma^2_{X_j}}{k(\tau,\alpha)^2\sigma^2_{X_j}} + \frac{\sigma^2_{X_k}}{k(\tau,\alpha)^2\sigma^2_{X_k}} + 2\frac{\sigma_{X_j\,X_k}}{k(\tau,\alpha)^2\sigma_{X_j}\sigma_{X_k}}\right).$$

In a more compact form

$$\sigma^2_{(j\,k)} = \frac{1}{k(\tau,\alpha)^2}\left(1 + \rho_{j\,k}\right).$$

By the affine invariance of the elliptical family, $\mathrm{IQR}^{(j\,k)\,\xi} = k(\xi,\alpha)\sigma_{(j\,k)}$. Substituting in $\sigma^2_{(j\,k)}$

$$\mathrm{IQR}^{(j\,k)\,\xi} = \frac{k(\xi,\alpha)}{k(\tau,\alpha)}\sqrt{1+\rho_{j\,k}} = s(\xi,\tau,\alpha)\sqrt{1+\rho_{j\,k}}.$$

In the Gaussian case $k(\tau,\alpha) = k(\tau)$ and $k(\xi,\alpha) = k(\xi)$. We normalize $\mathrm{IQR}^{\xi}_{(j\,k)}$ by $\frac{k(\tau)}{k(\xi)} = s_g(\xi,\tau)$ yielding

$$\mathrm{TailCoR}^{(j\,k)\,\xi} = s_g(\xi,\tau)s(\xi,\tau,\alpha)\sqrt{1+\rho_{j\,k}}.$$

The same proof follows for $\rho_{j\,k} < 0$ and $Z_t^{(j\,k)} = \frac{1}{\sqrt{2}}(Y_{j\,t} - Y_{k\,t})$, except that $\sqrt{1+\rho_{j\,k}}$ is replaced by $\sqrt{1-\rho_{j\,k}}$. This change is unsubstantial since both expressions are equal ($\rho_{j\,k}$ is positive in $\sqrt{1+\rho_{j\,k}}$ and negative in $\sqrt{1-\rho_{j\,k}}$). Hence

$$\mathrm{TailCoR}^{(j\,k)\,\xi} = s_g(\xi,\tau)s(\xi,\tau,\alpha)\sqrt{1+|\rho_{j\,k}|}.$$

Q.E.D.

Proof of Theorem 2

By **E1**, $\hat{\text{TailCoR}}_T^{(j\,k)\,\xi} = 2s_g(\xi,\tau)\hat{\text{Q}}_T^{(j\,k)\,\xi}$. The term $2s_g(\xi,\tau)$ is a deterministic scale shift and the only source of randomness is $\hat{\text{Q}}_T^{(j\,k)\,\xi}$. Hence

$$\begin{aligned} E(\hat{\text{TailCoR}}_T^{(j\,k)\,\xi}) &= 2s_g(\xi,\tau)E(\hat{\text{Q}}_T^{(j\,k)\,\xi}) \text{ and} \\ Var(\hat{\text{TailCoR}}_T^{(j\,k)\,\xi}) &= 4s_g(\xi,\tau)^2 Var(\hat{\text{Q}}_T^{(j\,k)\,\xi}). \end{aligned}$$

By the asymptotic properties of sample quantiles under S–mixing (Dominicy et al. (2013)) and the delta method the proof is completed. Q.E.D.

Proof of Theorem 3

It follows the same lines as the proof of Theorem 2. By **E1**, $vech\hat{\mathbf{TailCoR}}_{\tilde{\mathbf{Z}},T}^{\xi} = 2s_g(\xi,\tau)\hat{\mathbf{Q}}_T^{\xi}$, where $\hat{\mathbf{Q}}_T^{\xi}$ is a $\frac{\tilde{N}(\tilde{N}+1)}{2} \times 1$ vector of sample quantiles. By the asymptotic properties of vectors of sample quantiles under S–mixing (Dominicy et al. (2013)) and the delta method the proof is completed. Q.E.D.

Semi moments

We review standard definitions and properties for positive semi moments that we need. Results for negative semi moments are analogous. Let $I_{\{X_j>\mu_j\}}$ be an indicator function that takes value one if $X_j > \mu_j$. The positive semi variance of X_j is $\sigma_j^{+\,2} = E[((X_j-\mu_j)I_{\{X_j>\mu_j\}})^2]$. It possesses the standard location and scale shift property: let a and b be real and a positive real numbers respectively, then the positive semi variance of $a+bX_j$ is $b^2\sigma_j^{+\,2}$. The positive semi covariance between X_j and X_k is $\sigma_{j\,k}^{+} = E[((X_j-\mu_j)I_{\{X_j>\mu_j\}})((X_k-\mu_k)I_{\{X_k>\mu_k\}})]$. It is invariant to location shifts but not to re-scaling: if c is real and d is real positive, the positive semi covariance between $a+bX_j$ and $c+dX_j$ (a and b defined as above) is $bd\sigma_{j\,k}^{+}$. The positive semi correlation between X_j and X_k is $\rho_{j\,k}^{+} = \frac{\sigma_{j\,k}^{+}}{\sigma_j^{+}\sigma_k^{+}}$. Last, the positive semi variance of a sum, say $X_j + X_k$ is

$$\sigma_{X_j+X_k}^{+\,2} = E[((X_j-\mu_j)I_{\{X_j>\mu_j\}} + (X_k-\mu_k)I_{\{X_k>\mu_k\}})^2] = \sigma_j^{+\,2} + \sigma_k^{+\,2} + 2\sigma_{j\,k}^{+}.$$

Proof of Theorem 4

Since the proof follows the same lines as the proof of Theorem 1, we only show the main steps for the upside $\text{TailCoR}^{(j\,k)\,\xi\,+}$. The proof for $\text{TailCoR}^{(j\,k)\,\xi\,-}$ is equivalent.

We first consider the case when $\rho_{j\,k} > 0$ and hence $Z_t^{(j\,k)} = \frac{1}{\sqrt{2}}(Y_{j\,t} + Y_{k\,t})$. The positive semi variance of $Y_{j\,t}$ is

$$\sigma_{Y_j}^{+\,2} = \frac{\sigma_{X_j}^{+\,2}}{(\mathrm{IQR}_j^{\tau\,+})^2},$$

and likewise for $Y_{k\,t}$. The positive semi variance of $Z_t^{(j\,k)}$ is

$$\sigma_{(j\,k)}^2 = \frac{\sigma_{X_j}^{+\,2}}{(\mathrm{IQR}_j^{\tau\,+})^2}\sin^2\phi + \frac{\sigma_{X_k}^{+\,2}}{(\mathrm{IQR}_k^{\tau\,+})^2}\cos^2\phi + 2\sigma_{Y_j Y_k}^{+}\sin\phi\cos\phi,$$

where $\sigma_{Y_j Y_k}^{+}$ is the positive semi covariance between $Y_{j\,t}$ and $Y_{k\,t}$. Since $\mathrm{IQR}_j^{\tau\,+} = k(\tau,\alpha,\gamma)^+\sigma_{X_j}^{+}$, $\mathrm{IQR}_k^{\tau\,+} = k(\tau,\alpha,\gamma)^+\sigma_{X_k}^{+}$, $\sin^2\phi + \cos^2\phi = 1$, and by using the positive semi correlation, $\sigma_{(j\,k)}^{+\,2}$ simplifies to

$$\sigma_{(j\,k)}^{+\,2} = \frac{1}{k(\tau,\alpha,\gamma)^2}\left(1 + 2\rho_{j\,k}^{+}\sin\phi\cos\phi\right).$$

Substituting $\sigma_{(j\,k)}^{+}$ for $\mathrm{IQR}^{(j\,k)\,\xi\,+}/k(\xi,\alpha,\gamma)^+$

$$\mathrm{IQR}^{(j\,k)\,\xi\,+} = s(\xi,\tau,\alpha,\gamma)^+\sqrt{1 + 2\rho_{j\,k}^{+}\sin\phi\cos\phi},$$

where $s(\xi,\tau,\alpha,\gamma)^+ = k(\xi,\alpha,\gamma)^+/k(\tau,\alpha,\gamma)^+$. Multiplying by the normalization $2s_g(\xi,\tau)$ ends the proof.

The same proof follows for $\rho_{j\,k} < 0$ and $Z_t^{(j\,k)} = \frac{1}{\sqrt{2}}(Y_{j\,t} - Y_{k\,t})$, except that $\sqrt{1 + 2\rho_{j\,k}^{+}\sin\phi\cos\phi}$ is replaced by $\sqrt{1 - 2\rho_{j\,k}^{+}\sin\phi\cos\phi}$. This change is unsubstantial since both expressions are equal ($\rho_{j\,k}^{+}$ is positive in $\sqrt{1 + 2\rho_{j\,k}^{+}\sin\phi\cos\phi}$ and negative in $\sqrt{1 - 2\rho_{j\,k}^{+}\sin\phi\cos\phi}$). Hence

$$\mathrm{TailCoR}^{(j\,k)\,\xi} = 2s_g(\xi,\tau)s(\xi,\tau,\alpha,\gamma)^+\sqrt{1 + 2|\rho_{j\,k}^{+}|\sin\phi\cos\phi}.$$

Q.E.D.

Appendix R: additional empirical results

Table 1: Alternatives measures

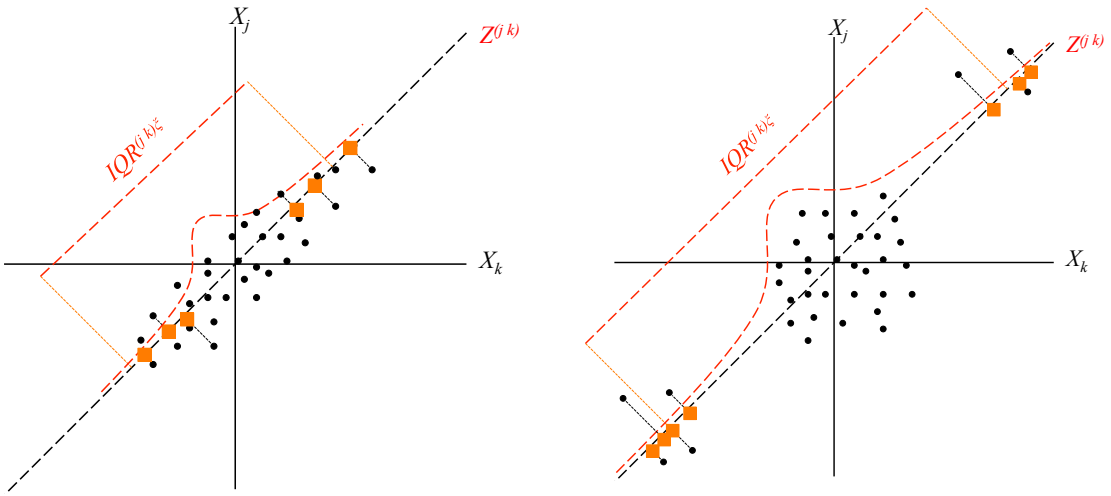
Downside exceedance correlation: $\theta^{-}(< 0.025)$								
	BAC	BK	C	GS	JPM	MS	WFC	STT
BAC	1.00	0.53	0.53	0.56	0.67	0.53	0.59	0.34
BK	0.53	1.00	0.66	0.36	0.61	0.47	0.71	0.64
C	0.53	0.66	1.00	0.63	0.16	0.34	0.49	0.78
GS	0.56	0.36	0.63	1.00	0.29	0.46	0.07	0.06
JPM	0.67	0.61	0.16	0.29	1.00	0.27	0.63	0.45
MS	0.53	0.47	0.34	0.46	0.27	1.00	0.23	0.02
WFC	0.59	0.71	0.49	0.07	0.63	0.23	1.00	0.42
STT	0.34	0.64	0.78	0.06	0.45	0.02	0.42	1.00
Upside exceedance correlation: $\theta^{+}(> 0.975)$								
	BAC	BK	C	GS	JPM	MS	WFC	STT
BAC	1.00	0.46	0.39	0.18	0.55	0.06	0.61	-0.06
BK	0.46	1.00	0.31	0.42	0.64	0.36	0.49	0.55
C	0.39	0.31	1.00	0.38	0.21	0.10	0.25	0.49
GS	0.18	0.42	0.38	1.00	0.15	0.45	0.23	0.57
JPM	0.55	0.64	0.21	0.15	1.00	0.17	0.65	0.16
MS	0.06	0.36	0.10	0.45	0.17	1.00	0.55	0.16
WFC	0.61	0.49	0.25	0.23	0.65	0.55	1.00	0.37
STT	-0.06	0.55	0.49	0.57	0.16	0.16	0.37	1.00
Copula-t tail dependence: τ_p								
	BAC	BK	C	GS	JPM	MS	WFC	STT
BAC	1.00	0.27	0.40	0.28	0.40	0.30	0.40	0.29
BK	0.27	1.00	0.28	0.27	0.33	0.28	0.30	0.38
C	0.40	0.28	1.00	0.31	0.39	0.32	0.27	0.28
GS	0.28	0.27	0.31	1.00	0.31	0.37	0.21	0.22
JPM	0.40	0.33	0.39	0.31	1.00	0.32	0.40	0.30
MS	0.30	0.28	0.32	0.37	0.32	1.00	0.25	0.26
WFC	0.40	0.30	0.27	0.21	0.40	0.25	1.00	0.29
STT	0.29	0.38	0.28	0.22	0.30	0.26	0.29	1.00
Non–parametric tail dependence: $\tau_{np} \times 100$								
	BAC	BK	C	GS	JPM	MS	WFC	STT
BAC	100	97.57	97.55	97.49	97.60	97.56	97.59	97.54
BK	97.57	100	97.55	97.47	97.53	97.51	97.65	97.52
C	97.55	97.55	100	97.63	97.62	97.47	97.50	97.59
GS	97.49	97.47	97.63	100	97.53	97.58	97.50	97.51
JPM	97.60	97.53	97.62	97.53	100	97.52	97.58	97.57
MS	97.56	97.51	97.47	97.58	97.52	100	97.58	97.60
WFC	97.59	97.65	97.50	97.50	97.58	97.58	100	97.50
STT	97.54	97.52	97.59	97.51	97.57	97.60	97.50	100

Tables and Figures

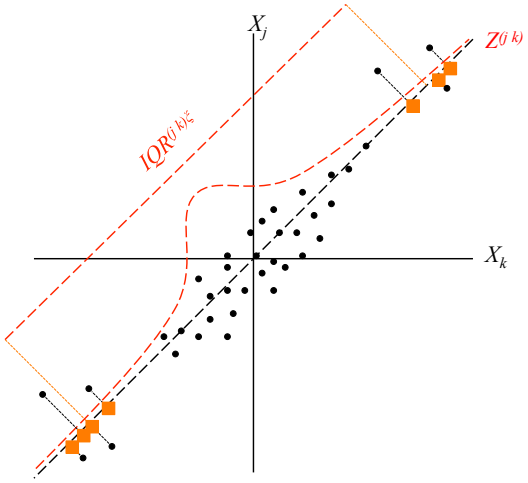
Figure 1: Diagrammatic representation of TailCoR

(a) Linear relation

(b) Non-linear relation

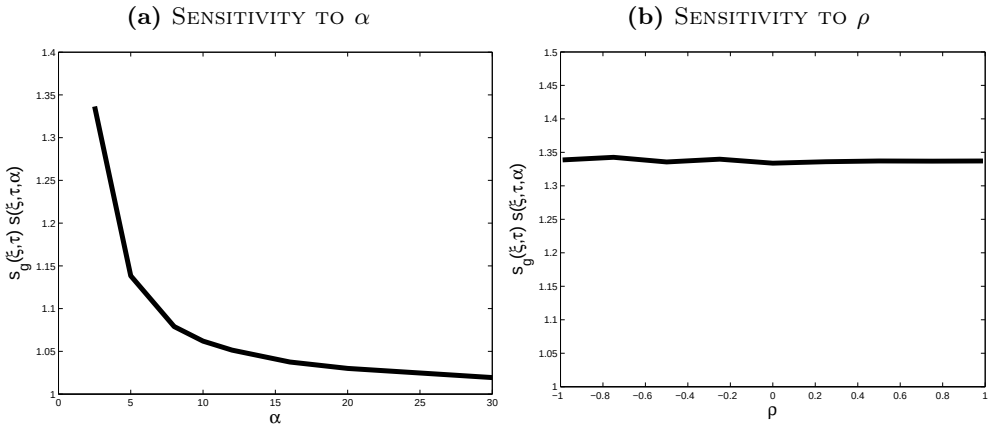


(c) Linear and non-linear relation



Scatter plots where X_j and X_k are positively related (the pairs are depicted with circles). Projecting the observations onto the 45-degree line produces the random variable $Z^{(j\,k)}$, depicted with squares. Because of representation purposes we show the projection only for the observations on the tails but the reader should keep in mind that the projection is done for all the observations. Panel (a) shows the case of only linear relation. Panel (b) displays the case of only non-linear relation. Panel (c) shows the scenario of both linear and non-linear relations.

Figure 2: Sensitivity of $s_g(\xi,\tau)s(\xi,\tau,\alpha)$ to α and ρ



Panel (a) shows the sensitivity of the non-linear correlation to α. The tail index varies from 2.5 to 30. Panel (b) shows the sensitivity to ρ (for $\alpha = 2.5$). Both plots are for $\tau = 0.75$ and $\xi = 0.95$.

Table 2: Mean and variances

T	H	Mean	Variance
100	100	1.67	0.27
100	500	1.68	0.24
1000	100	1.67	0.22
1000	500	1.64	0.08
5000	100	1.64	0.07
5000	500	1.64	0.04

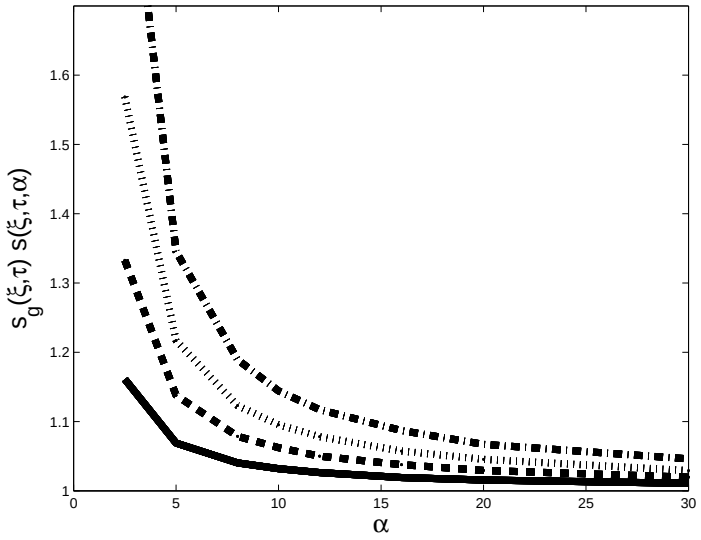
Mean and variances of the TailCoR estimates and for all the combinations of the sample size and replications.

Table 3: Descriptive statistics

Bank	Median	IQR_{75}	Kurtosis	Skewness
BAC	-0.02	1.15	0.58	-0.01
BK	-0.03	1.09	0.78	0.03
C	-0.02	1.15	0.58	-0.02
GS	-0.02	1.19	0.35	0.02
JPM	-0.02	1.18	0.44	0.04
MS	-0.03	1.17	0.44	0.01
WFC	-0.03	1.18	0.49	0.02
STT	-0.02	1.09	0.64	0.02

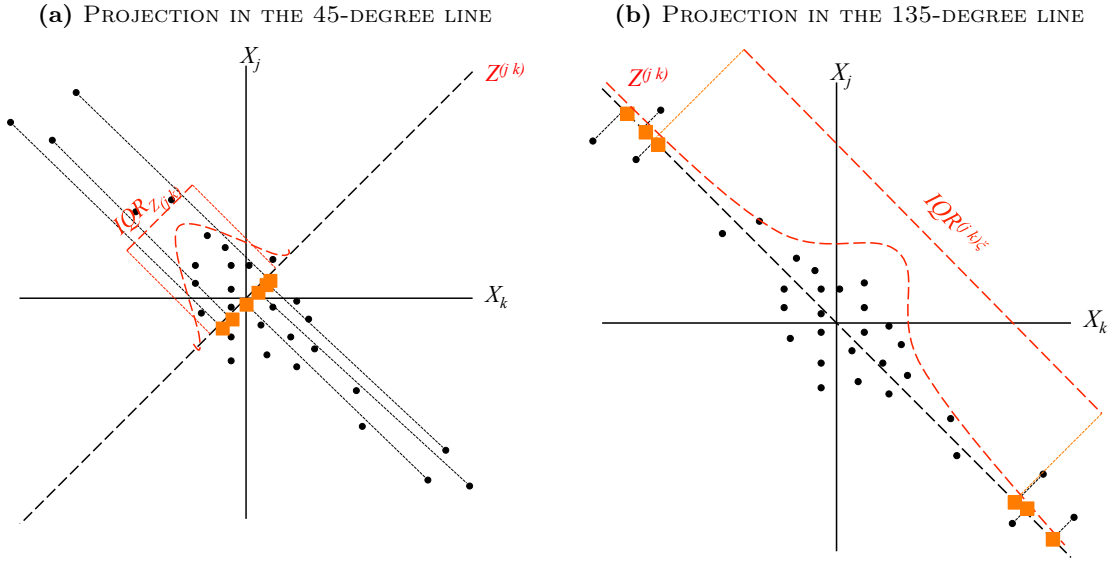
Kurtosis is computed as excess with respect to the Gaussian: $\text{IQR}^{0.975}/\text{IQR}^{0.75} - 2.91$. The skewness is computed as $(Q^{0.975} - Q^{0.50}) - (Q^{0.50} - Q^{0.025})$.

Figure 3: Sensitivity of $s_g(\xi,\tau)s(\xi,\tau,\alpha)$ to α and ξ



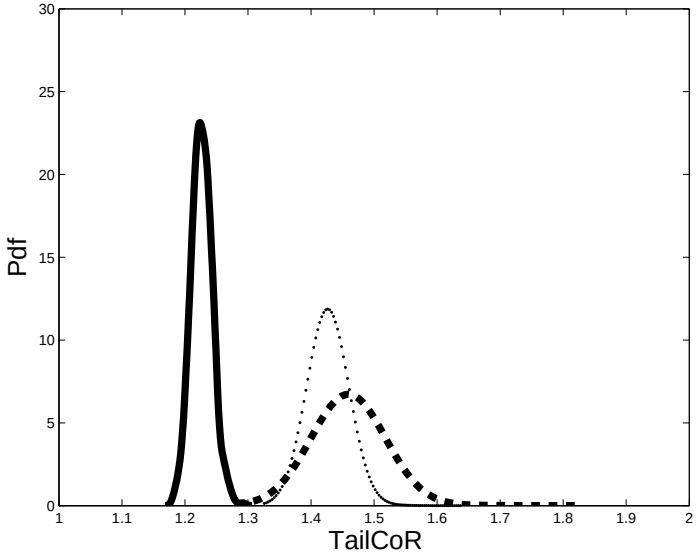
Sensitivity of the non-linear correlation to α for the Student-t distribution. The tail index varies from 2.5 to 30. Each line is for a value of ξ: 0.90 (solid line), 0.95 (thick dashes), 0.975 (thin dashes) and 0.99 (thick and thin dashes).

Figure 4: A diagrammatic representation of TailCoR for negative relation



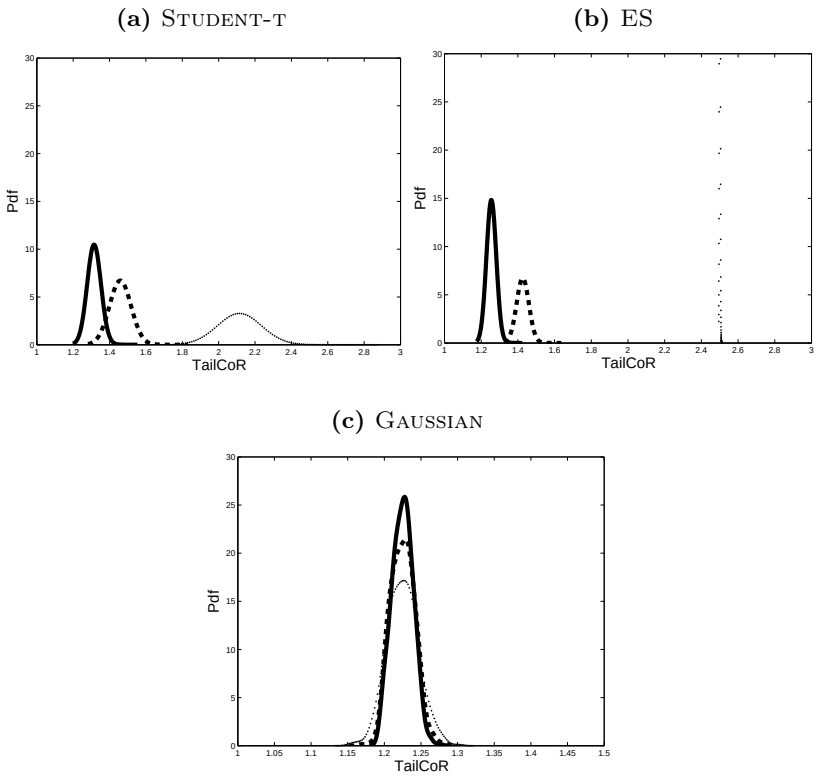
Scatter plots, along with the 45- (panel a) and 135-degree (panel b) lines, where X_j and X_k are negatively related.

Figure 5: TailCoR for different distributions



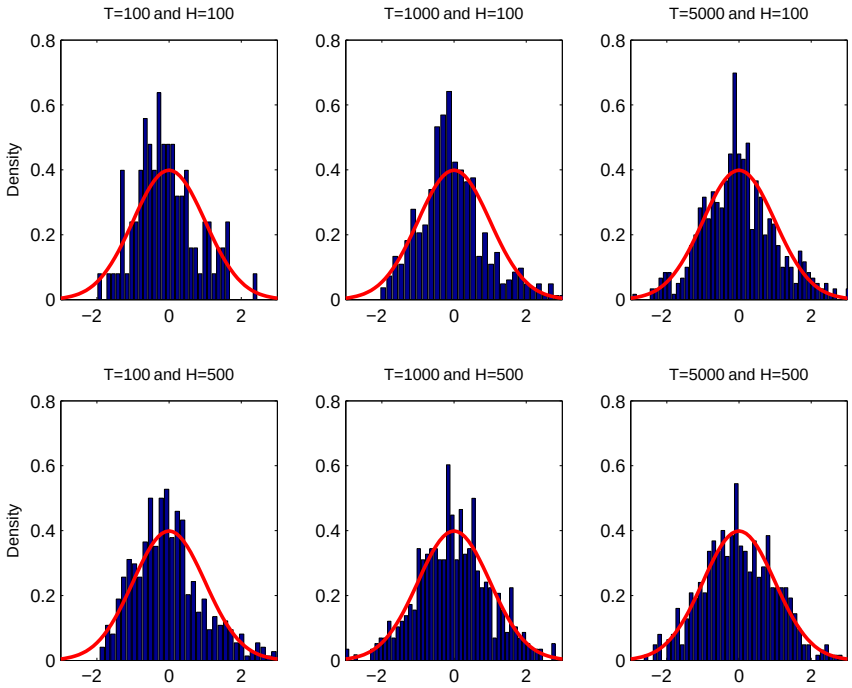
Distribution of 500 estimated TailCoR for $\xi = 0.95$ and three distributions: Student-t with $\alpha = 2.5$ (dashed line), ES with $\alpha = 1.5$ (dotted line), and Gaussian (solid line).

Figure 6: Sensitivity of TailCoR to ξ



Sensitivity of TailCoR to ξ for the Student-t with $\alpha = 2.5$ (panel a), ES with $\alpha = 1.5$ (panel b), and Gaussian (panel c). Each line is the density of the 500 estimates of TailCoR for different values of ξ: 0.90 (solid line), 0.95 (dashed) and 0.99 (dotted).

Figure 7: Convergence in distribution



Histograms of the standardized estimates against a standardized Gaussian distribution (red line) for all the combinations of the sample size and replications (indicated in the top of each plot).

Table 4: TailCoR and its contributions

	BAC	BK	C	GS	JPM	MS	WFC	STT
Matrix of TailCoR								
BAC	1.68	1.59	1.59	1.47	1.53	1.51	1.52	1.53
	(0.06)	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.04)
BK	1.59	1.78	1.59	1.51	1.53	1.54	1.54	1.64
	(0.04)	(0.06)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)
C	1.59	1.59	1.68	1.50	1.52	1.50	1.52	1.55
	(0.05)	(0.04)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
GS	1.47	1.51	1.50	1.57	1.46	1.53	1.43	1.49
	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
JPM	1.53	1.53	1.52	1.46	1.62	1.48	1.50	1.52
	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)
MS	1.51	1.54	1.50	1.53	1.48	1.61	1.48	1.54
	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)
WFC	1.52	1.54	1.52	1.43	1.50	1.48	1.64	1.51
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
STT	1.53	1.64	1.55	1.49	1.52	1.54	1.51	1.71
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)
Matrix of linear contributions $\sqrt{1+\rho}$								
	BAC	BK	C	GS	JPM	MS	WFC	STT
BAC	1.41	1.34	1.38	1.35	1.38	1.36	1.38	1.32
	(0)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)
BK	1.34	1.41	1.34	1.35	1.36	1.35	1.36	1.37
	(0.01)	(0)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)
C	1.38	1.34	1.41	1.35	1.38	1.37	1.37	1.32
	(0.00)	(0.01)	(0)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)
GS	1.35	1.35	1.35	1.41	1.37	1.39	1.35	1.34
	(0.01)	(0.01)	(0.01)	(0)	(0.00)	(0.00)	(0.01)	(0.01)
JPM	1.38	1.36	1.38	1.37	1.41	1.37	1.38	1.35
	(0.00)	(0.00)	(0.00)	(0.00)	(0)	(0.00)	(0.00)	(0.01)
MS	1.36	1.35	1.37	1.39	1.37	1.41	1.35	1.34
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0)	(0.01)	(0.01)
WFC	1.38	1.36	1.37	1.35	1.38	1.35	1.41	1.34
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0)	(0.01)
STT	1.32	1.37	1.32	1.34	1.35	1.34	1.34	1.41
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0)
Matrix of non-linear contributions $s_g(0.975, 0.75)s(0.975, 0.75, \alpha)$								
	BAC	BK	C	GS	JPM	MS	WFC	STT
BAC	1.19	1.19	1.15	1.09	1.10	1.11	1.10	1.16
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)
BK	1.19	1.26	1.18	1.12	1.12	1.13	1.14	1.20
	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.04)
C	1.15	1.18	1.19	1.11	1.10	1.10	1.11	1.17
	(0.04)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
GS	1.09	1.12	1.11	1.11	1.07	1.10	1.06	1.12
	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
JPM	1.10	1.12	1.10	1.07	1.14	1.07	1.09	1.12
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
MS	1.11	1.13	1.10	1.10	1.07	1.14	1.09	1.15
	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)
WFC	1.10	1.14	1.11	1.06	1.09	1.09	1.16	1.13
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
STT	1.16	1.20	1.17	1.12	1.12	1.15	1.13	1.21
	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)

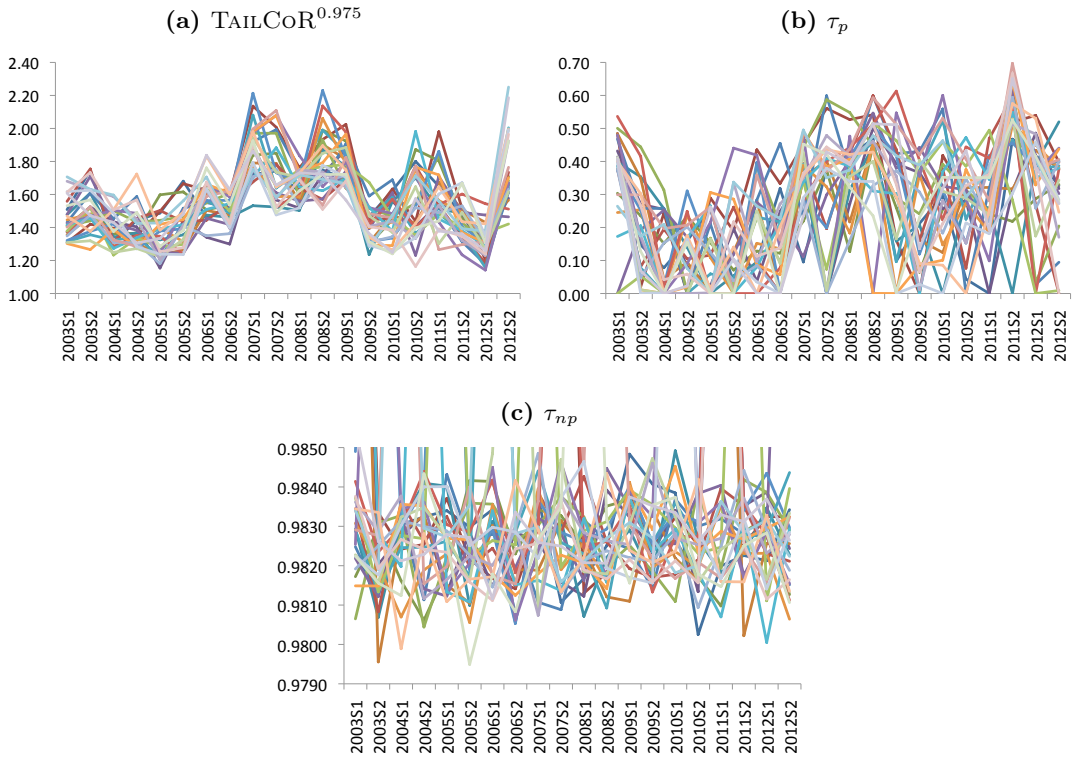
Matrix of estimated TailCoRs (upper panel) and their linear and non-linear components (middle and bottom panels respectively). Number in parenthesis are standard errors.

Table 5: Correlations between all measures

	$\text{TailCoR}^{0.975}$	$s(0.975, 0.75, \alpha)$	$\sqrt{1+\rho}$	$\theta^+(> 0.975)$	$\theta^-(< 0.025)$	τ_p	τ_{np}
$\text{TailCoR}^{0.975}$	1	0.89	0.52	0.74	0.72	0.75	0.72
$s(0.975, 0.75, \alpha)$	0.89	1	0.07	0.51	0.54	0.43	0.45
$\sqrt{1+\rho}$	0.52	0.07	1	0.66	0.57	0.83	0.73
$\theta^+(> 0.975)$	0.74	0.51	0.66	1	0.92	0.82	0.80
$\theta^-(< 0.025)$	0.72	0.54	0.57	0.92	1	0.74	0.70
τ_p	0.75	0.43	0.83	0.82	0.74	1	0.98
τ_{np}	0.70	0.43	0.71	0.78	0.69	0.97	1

Correlations between $\text{TailCoR}^{0.975}$, $\theta^+(> 0.975)$, $\theta^-(< 0.025)$, τ_p, and τ_{np}, as well as with the linear and non-linear components of TailCoR. The correlations are computed by vectorizing the fives matrices, and calculating the Pearson correlations between them.

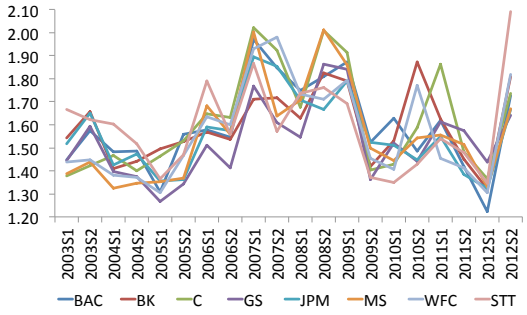
Figure 8: TailCoR and tail dependence coefficients



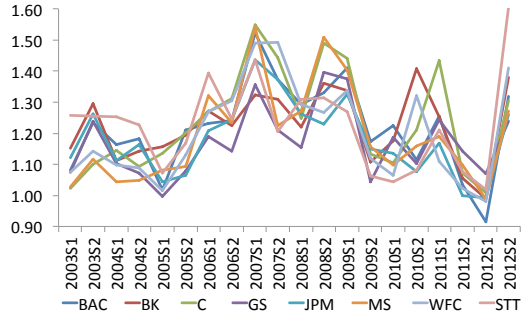
Panel (a) shows the evolution of TailCoR for all pairs of banks, while panels (b) and (c) show the parametric and non-parametric tail dependence coefficients.

Figure 9: TailCoR and its contributions

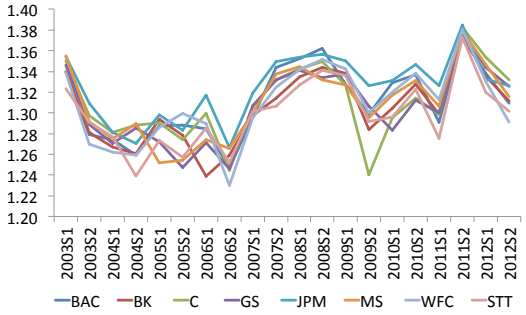
(a) TailCoR$^{0.975}$



(b) $s(0.975, 0.75)s(0.975, 0.75, \alpha)$



(c) $\sqrt{1+\rho}$



Panel (a) shows the evolution of TailCoR for all banks. Each line is the cross-sectional average of one bank with respect to the others. Panels (b) and (c) show the evolution of the non-linear and linear contributions respectively.