Quantitative Finance and Economics

http://www.aimspress.com/journal/QFE

QFE, 2(3): 590–614.
DOI:10.3934/QFE.2018.3.590
Received: 26 March 2018
Accepted: 24 May 2018
Published: 01 August 2018

Research article

How often is the financial market going to collapse?

Gabriel Frahm*

Helmut Schmidt University, Department of Mathematics and Statistics, Holstenhofweg 85, D–22043 Hamburg, Germany

* **Correspondence:** Email: frahm@hsu-hh.de; Tel: +49 (0)40 65412791; Fax: +49 (0)40 65412023.

Abstract: Copula theory is used to investigate the phenomenon of extremal dependence. An analytical expression for the extremal-dependence coefficient (EDC) of regularly varying elliptically distributed random vectors is derived. The EDC represents a natural measure of systemic risk. Extreme value theory is applied in order to estimate the systemic risk of the G–7 countries. The given results are quite sensitive to the tail index of asset returns and thus a scenario analysis is conducted. In the worst case, the probability that the entire market crashes during 10 years exceeds 50%. Hence, we must not neglect the risk of a financial collapse during a relatively short period of time.

Keywords: copula theory; extremal dependence; extreme value theory; ruin; tail dependence
JEL codes: G01, G15

1. Motivation

It is a stylized fact that the distribution of short-term asset returns exhibits heavy tails and tail dependence. These phenomena, together with other well-known stylized facts that can be observed for stocks, stock indices, foreign exchange rates, etc., are often reported in the literature during the last decades (see, e.g., Bouchaud et al., 1997; Breymann et al.; 2003; Cont, 2001; Ding et al., 1993; Dobri et al., 2013; Eberlein and Keller, 1995; Embrechts et al., 1997; Engle, 1982; Fama, 1965; Frahm and Jaekel, 2015; Junker and May, 2005; Mandelbrot, 1963; McNeil et al., 2005; Mikosch, 2003). This work proposes a measure for systemic risk. For this reason, I focus on the G–7 countries, i.e., Canada, France, Germany, Italy, Japan, UK, and USA. These countries represent the 7 largest economies worldwide. According to Credit Suisse's Global Wealth Databook (Shorrocks et al., 2017), they accumulate almost 63% of the global net wealth. Hence, we can expect that the G–7 countries have a major impact on the economic situation of each other country in the world.

Figure 1 shows the normal Q–Q plot of daily log-returns on the MSCI country index for Canada from 1999-01-04 to 2018-03-02, which means that the chosen period covers the dot-com collapse at the beginning of 2000 and the financial crisis 2007–2008. The index points are based on USD total returns and the number of observations is $n = 4804$. Figure 10 in the appendix completes this picture by referring to France, Germany, Italy, Japan, UK, and USA. The QQ-plots reveal that the probability of extremes is much higher than suggested by the normal distribution. Hence, the normal-distribution hypothesis is highly misleading—at least if we refer to daily log-returns. Nowadays, this simple but, nonetheless, far-reaching statement has become folklore in the finance literature.

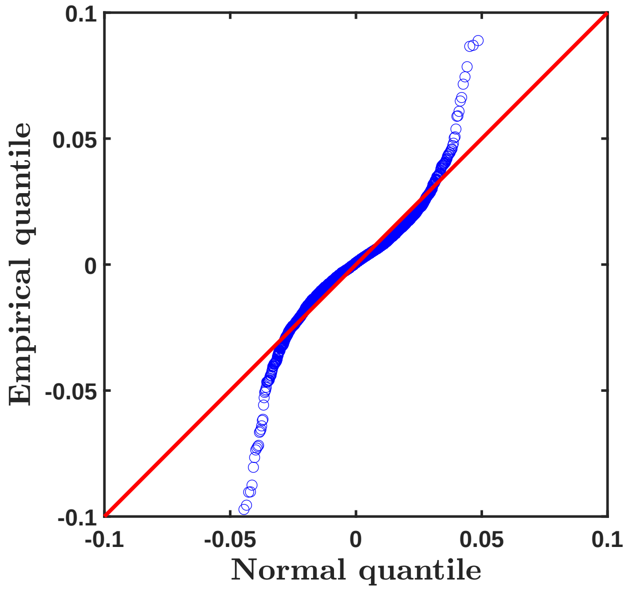

Figure 1. Normal Q–Q plot of daily log-returns from 1999-01-04 to 2018-03-02 on the MSCI country index for Canada.

Which model is appropriate if we aim at taking heavy tails and tail dependence properly into account? Figure 2 illustrates the joint distribution of the daily log-returns of Canada and France. The scatter plot reveals the following effects:

1. The central region of the distribution seems to be normal or, at least, elliptically contoured,
2. there is a large number of outliers or extreme values,
3. extreme values typically occur simultaneously, and
4. their distribution is asymmetric.

The last point is based on the observation that the magnitude of extreme values on the lower left appears to be larger compared to the upper right of the scatter plot. The same effects can typically be observed when comparing all G–7 countries with each other (see Figure 11, which can be found in the appendix).

Let F be the (cumulative) distribution function of some random variable, whereas F^{-1} denotes the associated quantile function, i.e., $p \mapsto F^{-1}(p) := \inf\{x \in \mathbb{R} : F(x) \geq p\}$. More specifically, let $X = (X_1, X_2, \ldots, X_7)$ be the vector of daily log-returns on the MSCI indices for the G–7 countries. Further, let F_i be the distribution function of the log-return on Country $i = 1, 2, \ldots, 7$. The event $X_i \leq F_i^{-1}(p)$ with $p \in (0, 1)$ is said to be a p-shortfall, where p is the corresponding shortfall probability.* The

*If F_i is strictly increasing, then $-F_i^{-1}(p)$ is the value at risk of X_i at the confidence level $1 - p$ (Artzner et al., 1999).

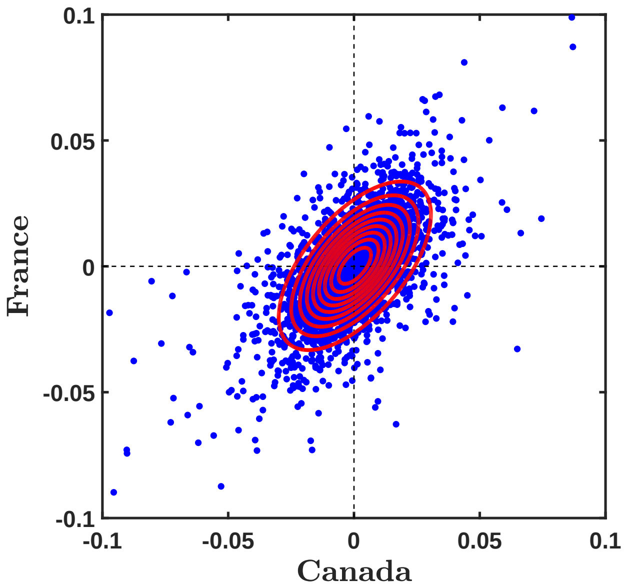

Figure 2. Scatter plot of daily log-returns from 1999-01-04 to 2018-03-02 on the MSCI country indices for Canada and France (blue points). The red contours represent the deciles of the bivariate normal distribution that is fitted to the data.

expected number of shortfalls during $m \in \mathbb{N}$ trading days amounts to mp and thus, on average, a p-shortfall occurs after p^{-1} trading days. Hence, the event $X_i \leq F_i^{-1}(p)$ is said to be a p^{-1}-day shortfall and, correspondingly, $F_i^{-1}(p)$ represents a p^{-1}-day quantile.

From time to time the financial market collapses. Figure 3 shows the historical log-performance of each G–7 country index from 1999-01-03 to 2018-03-02. The indices are normed at the beginning of the observation period. It reveals that the G–7 countries are typically affected by the same economic shocks. However, there are a few exceptions. For example, Italy had a drawdown in 1999, while the other countries performed well during this period. Moreover, the recession in Europe from mid 2014 to the end of 2015 cannot be observed neither in Japan nor in the USA.

Figure 4 contains the number of G–7 countries that had a 200-day shortfall at the same trading day. We can see the financial turmoil after the dot-com bubble at the end of the 20th century and during the financial crisis 2007–2008. On November 6, 2008, the International Monetary Fund predicted a worldwide decrease of the gross domestic product for the developed economies and, concomitantly, all G–7 countries crashed during this trading day. Additionally, there are some simultaneous shortfalls in 2011, which occurred due to the Greek debt crisis.

Simultaneous shortfalls are not evenly spread over time. It is obvious that the systemic risk prevails in times of crisis, i.e., the probability of concomitant shortfalls substantially increases after the financial market collapses. Put another way, simultaneous shortfalls appear in clusters. In this work, I ignore the time-series aspect of simultaneous shortfalls and focus on the cross-sectional dependence structure of extreme asset returns. This can be done by means of extreme value theory. The basic methodology is presented in the next section.

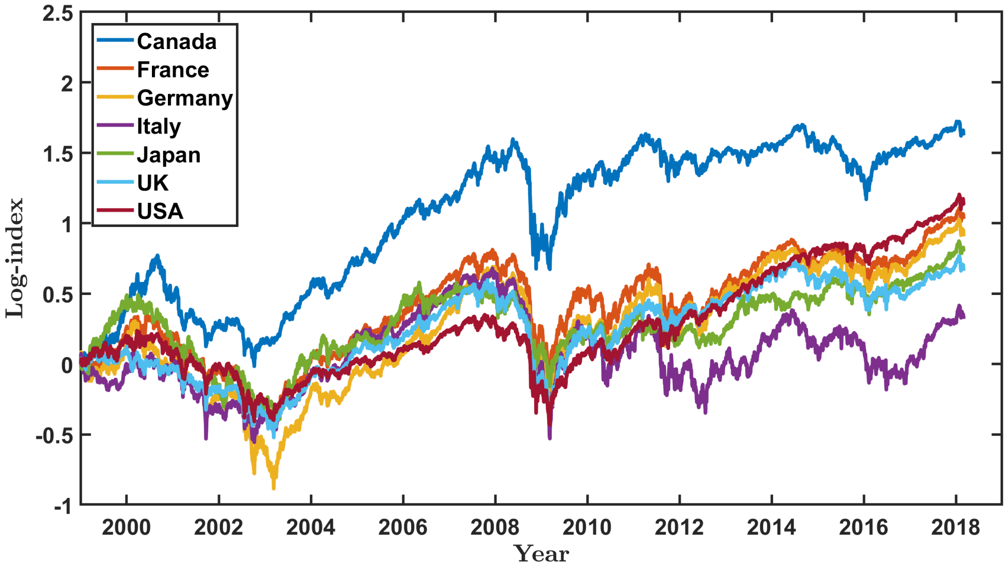

Figure 3. Natural logarithm of the G–7 country indices from 1999-01-03 to 2018-03-02.

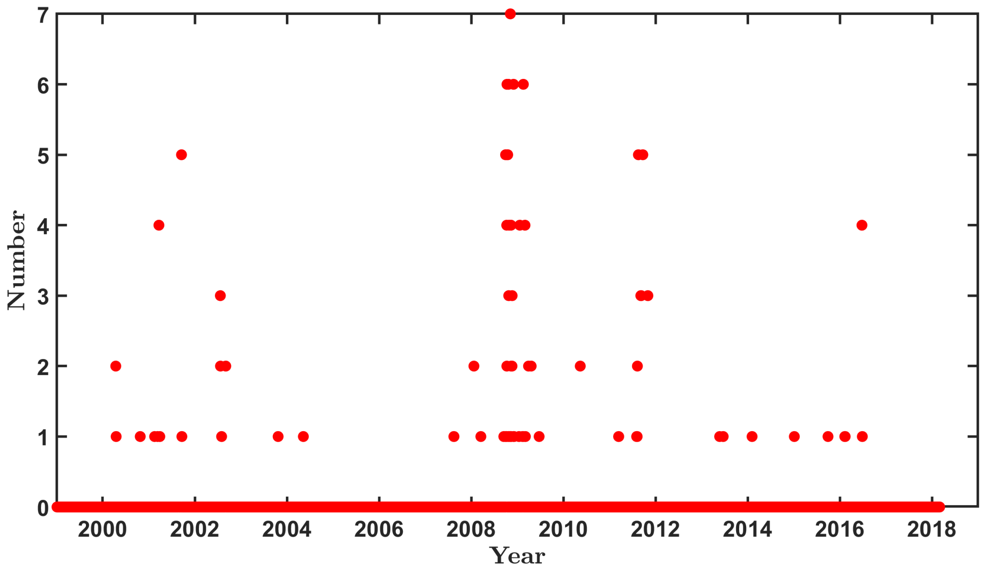

Figure 4. Number of G–7 countries that had a 200-day shortfall at the same trading day from 1999-01-04 to 2018-03-02.

2. Theoretical background

The phenomenon that extreme asset returns typically occur simultaneously is referred to as tail dependence, which is part of copula theory and extreme value theory. The reader can find a profound treatment of copula theory in Joe (1997) and Nelsen (2006), whereas Mikosch (2003) gives a nice overview of extreme value theory. I recapitulate some basic tools of copula theory and of extreme value theory in this section.

2.1. Tail vs. extremal dependence

The reader needs no specific knowledge about copulas to understand the following arguments, but at least he should be aware of Sklar's theorem (Sklar, 1959): Let F be the joint distribution function of any random vector $X = (X_1, X_2, \dots, X_d)$ and F_i the (marginal) distribution function of X_i ($i = 1, 2, \dots, d$). Then there exists a distribution function $C\colon [0,1]^d \to [0,1]$ such that

$$F(x) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d))$$

for all $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$. The function C is referred to as the copula of X. It represents the joint distribution function of the random vector $U = (U_1, U_2, \dots, U_d)$ with $U_i := F_i(X_i)$. If the marginal distribution functions of X are continuous, each component of U is uniformly distributed on $[0,1]$ and C is unique on $[0,1]^d$. I maintain this assumption throughout this work.

The lower tail-dependence coefficient (TDC) of a pair of random variables X_i and X_j (Joe, 1997) is defined as

$$\lambda_{\mathrm{L}} := \lim_{p \searrow 0} \mathbb{P}(U_j \le p \,|\, U_i \le p) = \lim_{p \searrow 0} \frac{C_{ij}(p,p)}{p},$$

where C_{ij} is the copula of (X_i, X_j). It is implicitly assumed that the given limit exists. Correspondingly, the upper TDC is defined as

$$\lambda_{\mathrm{U}} := \lim_{p \nearrow 1} \mathbb{P}(U_j > p \,|\, U_i > p) = \lim_{p \nearrow 1} \frac{1 - 2p + C_{ij}(p,p)}{1-p}.$$

Loosely speaking, the lower TDC is the probability that Country j crashes given that Country i crashes or, equivalently, that Country i crashes given that Country j crashes. If λ_{L} or λ_{U} is positive, then X_i and X_j are said to be (lower or upper) tail dependent.

The TDC is a very popular risk measure. However, it is defined only for the bivariate case and so, when applying this measure, we must restrict to some pair of G–7 countries. There are several ways to extend the concept of tail dependence to the case of $d > 2$ (De Luca and Rivieccio, 2012; Ferreira and Ferreira, 2012). In this work, I focus on the notion of extremal dependence (Frahm, 2006). The extremal-dependence coefficient (EDC) introduced by Frahm (2006) seems to be a natural measure of systemic risk, i.e., the risk of a collapse of the *entire* financial market.

In the following, I use the shorthand notation

$$\min \zeta := \min\left\{\zeta_1, \zeta_2, \dots, \zeta_d\right\} \qquad \text{and} \qquad \max \zeta := \max\left\{\zeta_1, \zeta_2, \dots, \zeta_d\right\},$$

where $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_d)$ is any random vector.

Definition 1 (Lower and upper EDC)**.** *The lower EDC of X is defined as*

$$\varepsilon_{\mathrm{L}} := \lim_{p \searrow 0} \mathbb{P}(\max U \le p \,|\, \min U \le p),$$

whereas its upper EDC is defined as

$$\varepsilon_{\mathrm{U}} := \lim_{p \nearrow 1} \mathbb{P}(\min U > p \,|\, \max U > p).$$

We can also write, equivalently,

$$\varepsilon_{\mathrm{L}} = \lim_{p\searrow 0} \frac{\mathbb{P}(\max U \le p)}{\mathbb{P}(\min U \le p)} \qquad \text{and} \qquad \varepsilon_{\mathrm{U}} = \lim_{p\nearrow 1} \frac{\mathbb{P}(\min U > p)}{\mathbb{P}(\max U > p)}.$$

Hence, the lower EDC can be considered the probability that the entire market collapses given that some part of the market crashes. Put another way, it is the probability that the whole system breaks down if some part of the system fails. When this happens, the fundamental principle of modern portfolio theory (Markowitz, 1952), i.e., diversification, does no longer work. In this case, even *international* diversification (Jorion, 1985) does not help much.

Whenever ε_{L} or ε_{U} is positive, the components of X are said to be (lower or upper) extremal dependent. It can be shown that

$$\varepsilon_{\mathrm{L}} = \lim_{p\searrow 0} \frac{C(p,p,\ldots,p)}{1-\overline{C}(1-p,1-p,\ldots,1-p)} \qquad \text{and} \qquad \varepsilon_{\mathrm{U}} = \lim_{p\nearrow 1} \frac{\overline{C}(1-p,1-p,\ldots,1-p)}{1-C(p,p,\ldots,p)},$$

where $\overline{C}$ is the survival copula associated with C (Frahm, 2006). This is defined by

$$u \longmapsto \overline{C}(u) := \sum_{I\subseteq M} (-1)^{|I|} C\left((1-u_1)^{\mathbf{1}_{1\in I}}, (1-u_2)^{\mathbf{1}_{2\in I}}, \ldots, (1-u_d)^{\mathbf{1}_{d\in I}}\right),$$

where $u = (u_1, u_2, \ldots, u_d) \in [0,1]^d$, $M := \{1,2,\ldots,d\}$, and $\mathbf{1}$ denotes the indicator function. Since the marginal distribution functions of X are continuous, $\overline{C}$ represents the copula of $-X$.

For convenience, I recapitulate some basic results concerning the TDC and the EDC, which can be found in Frahm (2006).

Proposition 1. *Let λ_{L} and λ_{U} be the tail-dependence coefficients of any pair of random variables. Further, let ε_{L} and ε_{U} be the corresponding extremal-dependence coefficients. Then we have that*

$$\varepsilon_{\mathrm{L}} = \frac{\lambda_{\mathrm{L}}}{2-\lambda_{\mathrm{L}}} \qquad \textit{and} \qquad \varepsilon_{\mathrm{U}} = \frac{\lambda_{\mathrm{U}}}{2-\lambda_{\mathrm{U}}}.$$

Hence, the EDC is a convex function of the TDC and we have that $\varepsilon_{\mathrm{L}} < \lambda_{\mathrm{L}}$ for all $0 < \lambda_{\mathrm{L}} < 1$ as well as $\varepsilon_{\mathrm{U}} < \lambda_{\mathrm{U}}$ for all $0 < \lambda_{\mathrm{U}} < 1$ (see Figure 5).

Proposition 2. *Let X be a d-dimensional random vector with $d > 1$ and X_s any subvector of X. Further, let $\varepsilon_{\mathrm{L}}(X)$ be the lower EDC of X and $\varepsilon_{\mathrm{L}}(X_s)$ be the lower EDC of X_s. Similarly, let $\varepsilon_{\mathrm{U}}(X)$ be the upper EDC of X and $\varepsilon_{\mathrm{U}}(X_s)$ be the upper EDC of X_s. Then we have that*

$$\varepsilon_{\mathrm{L}}(X) \le \varepsilon_{\mathrm{L}}(X_s) \qquad \textit{and} \qquad \varepsilon_{\mathrm{U}}(X) \le \varepsilon_{\mathrm{U}}(X_s).$$

This means that if we extend our economy by adding some country, the extremal dependence cannot increase. In general, it decreases after an extension of the market because the greater the number of countries, the more unlikely it is that the entire world collapses, given that (at least) one country crashes. Nonetheless, we should keep in mind that the probability that some country falls into the abyss usually increases the larger the economy.

Proposition 1 and Proposition 2 imply that

$$\varepsilon_{\mathrm{L}}(X) \le \varepsilon_{\mathrm{L}}(X_i, X_j) = \frac{\lambda_{\mathrm{L}}(X_i, X_j)}{2-\lambda_{\mathrm{L}}(X_i, X_j)}$$

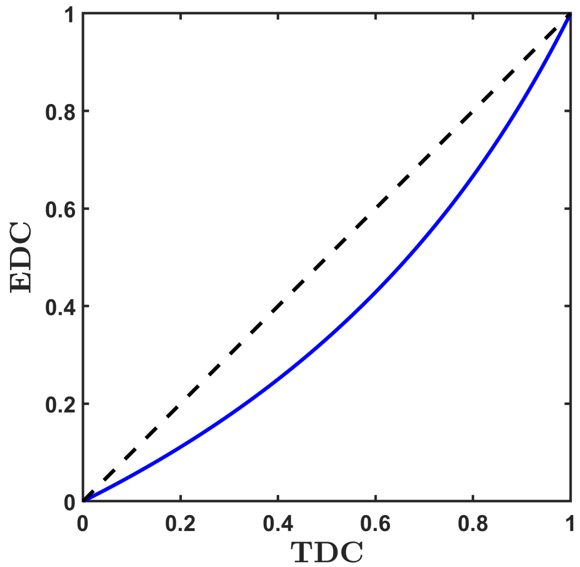

Figure 5. EDC of a 2-dimensional random vector as a function of the TDC.

for $i, j = 1, 2, \dots, d$, where $\lambda_L(X_i, X_j)$ denotes the TDC of X_i and X_j. Hence, the lower EDC of every 2-dimensional subvector of X is positive whenever the lower EDC of X is positive, which means that the lower TDC of each 2-dimensional subvector must be positive, too. Conversely, if the TDC of any subvector (X_i, X_j) is zero, the components of X cannot be extremal dependent. The same arguments apply to the upper risk measures. To sum up, if the components of a d-dimensional random vector $X = (X_1, X_2, \dots, X_d)$ are extremal dependent, then X_i and X_j are tail dependent for $i, j = 1, 2, \dots, d$, but if X_i and X_j are not tail dependent for any $i, j \in \{1, 2, \dots, d\}$, then neither the components of X can be extremal dependent.

Let the copula C be symmetric in the sense that $C(u) = \overline{C}(u)$ for all $u \in [0, 1]^d$. This sort of symmetry shall be referred to as transpositional symmetry. If C is transpositionally symmetric, its "lower left corner" coincides with its "upper right corner," in which case the lower EDC of X corresponds to its upper EDC. Then we can simply write $\varepsilon \equiv \varepsilon_L = \varepsilon_U$. A d-dimensional random vector X has a transpositionally symmetric copula if the distribution of X is symmetric, i.e., if there exists a location vector $\mu \in \mathbb{R}^d$ such that $X - \mu$ has the same distribution as $-(X - \mu)$.

The components of a random vector $X = (X_1, X_2, \dots, X_d)$ are said to be comonotonic if their dependence is perfectly positive. More precisely, $X_1, X_2, \dots, X_d$ are comonotonic if and only if there exist a random variable V and d strictly increasing functions of the form $f_i : \mathbb{R} \to \mathbb{R}$ such that $X_i = f_i(V)$ for $i = 1, 2, \dots, d$. In this case, the copula of X corresponds to the "minimum copula" $u \mapsto \min u$, which is called Fréchet-Hoeffding upper bound (Nelsen, 2006). Then both the lower and the upper EDC of X equal 1. By contrast, if the components of X are mutually independent, the copula of X corresponds to the "product copula" $u \mapsto \prod_{i=1}^{d} u_i$, in which case both the lower EDC and the upper EDC of X equal 0. Finally, if the dependence between two components X_i and X_j is perfectly negative, they are said to be countermonotonic. More precisely, X_i and X_j are countermonotonic if and only if there exist a random variable V, a strictly increasing function $f : \mathbb{R} \to \mathbb{R}$, and a strictly decreasing function $g : \mathbb{R} \to \mathbb{R}$ such that $X_i = f(V)$ and $X_j = g(V)$. The copula of X_i and X_j corresponds to the Fréchet-Hoeffding lower bound $(u_i, u_j) \mapsto \max\{u_i + u_j - 1, 0\}$ (Nelsen, 2006, p. 11). Then both the lower EDC and the upper EDC of (X_i, X_j) equal 0. Hence, in the bivariate case, the EDC does not distinguish between countermonotonicity and independence.

2.2. Regular variation of elliptical distributions

It is well-known that the multivariate normal distribution does not allow for tail dependence. Put another way, in this case we have that $\lambda_{\mathrm{L}} = \lambda_{\mathrm{U}} = 0$. This implies that the components of a normally distributed random vector cannot be extremal dependent either. In the risk-management literature, the class of elliptical distributions (Cambanis et al., 1981; Fang et al., 1990; Kelker, 1970) is often proposed as an appropriate alternative to the normal distribution (see, e.g., Bingham and Kiesel, 2002; Eberlein and Keller, 1995; Frahm, 2004; McNeil et al., 2005). Here, I will adopt this approach. Elliptical distributions cover the first three observations made in Figure 2, which are discussed in Section 1, and they are tractable even if the number of dimensions is high. The fourth phenomenon, namely that the distribution of extreme asset returns is asymmetric, cannot be explained by elliptical distributions. For this purpose, we could make use of generalized elliptical distributions (Frahm and Jaekel, 2015; Frahm, 2004), but this goes beyond the scope of this work.

A d-dimensional random vector X is said to be elliptically distributed if and only if there exist a vector $\mu \in \mathbb{R}^d$, a matrix $\Lambda \in \mathbb{R}^{d\times k}$, a nonnegative random variable $\mathcal{R}$, and a k-dimensional random vector S that is stochastically independent of $\mathcal{R}$ and uniformly distributed on the unit hypersphere $\{s \in \mathbb{R}^k : \|s\|_2 = 1\}$ such that $X = \mu + \Lambda\mathcal{R}S$ (Cambanis et al., 1981). The parameter μ is called the location vector, $\Sigma := \Lambda\Lambda'$ is referred to as the dispersion matrix, and $\mathcal{R}$ is said to be the generating variate of X. If $\Lambda_1\Lambda_1' = \Lambda_2\Lambda_2'$ for any $\Lambda_1, \Lambda_2 \in \mathbb{R}^{d\times k}$, then the random vectors $\Lambda_1 S$ and $\Lambda_2 S$ have the same distribution. That is, the distribution of X depends on Λ only through Σ and, without loss of generality, I assume that $\operatorname{rk}\Sigma = d = k$. The second moments of X are finite if and only if $\mathbf{E}(\mathcal{R}^2) < \infty$, in which case we have that $\mathbf{Var}(X) = \mathbf{E}(\mathcal{R}^2)\,\Sigma/d$. However, the dispersion matrix Σ exists (and is finite) even if $\mathbf{E}(\mathcal{R}^2) = \infty$. The distribution of X is symmetric around μ and so the lower EDC coincides with the upper EDC of X.

In general, the components of an elliptically distributed random vector X exhibit two sorts of dependencies, viz.

(1) linear dependence, which can be expressed by the dispersion matrix Σ and
(2) nonlinear dependence, which is determined by the generating variate $\mathcal{R}$.

For example, the (spherically distributed) random vector $X = S$ contains no linear dependence at all. Nonetheless, in a nonlinear manner, the components of X highly depend on each other because the generating variate $\mathcal{R} = 1$ forces them to be such that $\|X\|_2 = 1$. It is well-known that the components of X are mutually independent if and only if X possesses a normal distribution, i.e., $\mathcal{R}^2 = \chi^2_d$, and the off-diagonal elements of Σ are zero, i.e., the components of X are uncorrelated.

In risk management it is typically assumed that the survival function of $\mathcal{R}$ is regularly varying (Mikosch, 2003). This means that

$$\mathbb{P}(\mathcal{R} > r) = f(r)\, r^{-\alpha}, \qquad \alpha \geq 0,$$

for all $r > 0$, where f is a slowly varying function, i.e., $f(tr)/f(r) \to 1$ as $r \to \infty$ for all $t > 0$.[†] Put

[†]This implies that there exists some threshold $\tau > 0$ such that $f(r) > 0$ for all $r \geq \tau$.

another way, $r \mapsto \mathbb{P}(\mathcal{R} > r)$ tends to a power law. This is equivalent to

$$\frac{\mathbb{P}(\mathcal{R} > tr)}{\mathbb{P}(\mathcal{R} > r)} \longrightarrow t^{-\alpha}, \qquad r \longrightarrow \infty,$$

for all $t > 0$. In this case, the distribution of $\mathcal{R}$ is said to be heavy tailed and α represents its tail index. The lower α the heavier the tail of the distribution of $\mathcal{R}$. To keep the terminology simple, I will say that $\mathcal{R}$ itself is heavy tailed or, equivalently, regularly varying (with tail index α).

Further, a d-dimensional random vector X is said to be regularly varying with tail index $\alpha \geq 0$ if and only if there exists a d-dimensional random vector S that is distributed on the unit hypersphere $\mathcal{S}^{d-1} = \{s \in \mathbb{R}^d : \|s\| = 1\}$ such that

$$\frac{\mathbb{P}(\|X\| > tr, X/\|X\| \in B)}{\mathbb{P}(\|X\| > r)} \longrightarrow t^{-\alpha}\, \mathbb{P}(S \in B), \qquad r \longrightarrow \infty\,,$$

for all $t > 0$ and every Borel set $B \subseteq \mathcal{S}^{d-1}$ with $\mathbb{P}(S \in \partial B) = 0$ (Mikosch, 2003).[‡] Here, we can choose any arbitrary norm $\|\cdot\|$, but the unit hypersphere $\mathcal{S}^{d-1}$ depends on the choice of $\|\cdot\|$. However, the norm does not affect the tail index α (Hult and Lindskog, 2002).

Regular variation properties of elliptical distributions are investigated by Frahm (2006), Hult and Lindskog (2002) as well as Schmidt (2002). The latter authors focus on the relationship between regular variation and the TDC. By contrast, Frahm (2006) studies the EDC of regularly varying elliptically distributed random vectors. Suppose that X is elliptically distributed with location vector $\mu = 0$ and let $\|\cdot\|_\Sigma$ be the Mahalanobis norm, i.e., $\|x\|_\Sigma^2 = x'\Sigma^{-1}x$ for all $x \in \mathbb{R}^d$. Then we have that $\|X\|_\Sigma = \mathcal{R}$ and $X/\|X\|_\Sigma = \Lambda S$, which leads to

$$\frac{\mathbb{P}(\|X\|_\Sigma > tr, X/\|X\|_\Sigma \in B)}{\mathbb{P}(\|X\|_\Sigma > r)} = \frac{\mathbb{P}(\mathcal{R} > tr)}{\mathbb{P}(\mathcal{R} > r)} \cdot \mathbb{P}(\Lambda S \in B) \longrightarrow t^{-\alpha}\, \mathbb{P}(\Lambda S \in B),$$

where $S \in \{s \in \mathbb{R}^d : \|s\|_2 = 1\}$ and thus $\Lambda S \in \{s \in \mathbb{R}^d : \|s\|_\Sigma = 1\}$. That is, if the generating variate $\mathcal{R}$ is regularly varying, the random vector X inherits the tail index of $\mathcal{R}$. Moreover, the regular variation property is not affected by translations of X, i.e., the previous result holds true if $\mu \neq 0$ (Hult and Lindskog, 2002).

Now, suppose that $\mathcal{R}$ is regularly varying with tail index α and define

$$\Sigma =: \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1d} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \cdots & \sigma_d^2 \end{bmatrix}, \quad \sigma := \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_d \end{bmatrix}, \quad \rho := \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1d} \\ \rho_{21} & 1 & \cdots & \rho_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{d1} & \rho_{d2} & \cdots & 1 \end{bmatrix},$$

where $\rho_{ij} := \sigma_{ij}/(\sigma_i\sigma_j)$ for $i, j = 1, 2, \ldots, d$ with $\sigma_{ii} \equiv \sigma_i^2$ for $i = 1, 2, \ldots, d$. Thus, ρ represents the correlation matrix of X and we have that $\Sigma = \sigma\rho\sigma$. Since $\mathbf{E}(\mathcal{R}^\gamma) < \infty$ for all $\gamma < \alpha$ but $\mathbf{E}(\mathcal{R}^\gamma) = \infty$ for all $\gamma > \alpha$ (Embrechts et al., 1997), the second moment of $\mathcal{R}$ is infinite if its tail index is lower than 2. Then the covariance matrix of X remains undefined. However, ρ still exists and can be considered a "pseudo-correlation matrix" (Frahm, 2006).

[‡]Here, "∂B" denotes the boundary of B.

The parameters μ and σ affect only the marginal distribution functions of X but not its copula. For this reason, we may concentrate on the correlation matrix ρ and the distribution of $\mathcal{R}$ in order to calculate the EDC of X. The dispersion matrix Σ is positive definite and so the same holds true for ρ. We can choose any matrix $\sqrt{\rho} \in \mathbb{R}^{d\times d}$ with $\operatorname{rk}\sqrt{\rho} = d$ such that $\rho = \sqrt{\rho}\sqrt{\rho}'$ and thus $\Lambda = \sigma\sqrt{\rho}$. Now, define the random variables $Y := \min \sqrt{\rho}\,S$ and $Z := \max \sqrt{\rho}\,S$.

The following theorem represents the main theoretical result of this work.

Theorem 1. *Let X be a d-dimensional regularly varying elliptically distributed random vector with positive definite correlation matrix ρ and tail index* $\alpha \geq 0$. *Then both the lower and the upper EDC of X correspond to*

$$\varepsilon = \frac{\int_0^\infty y^\alpha \, dF_Y(y)}{\int_0^\infty z^\alpha \, dF_Z(z)},$$

where F_Y *and* F_Z *denote the distribution functions of* $Y = \min \sqrt{\rho}\,S$ *and* $Z = \max \sqrt{\rho}\,S$, *respectively.*

Proof. The copula of $X = \mu + \sigma\sqrt{\rho}\,\mathcal{R}S$ neither depends on μ nor on σ. Hence, we may focus on the standardized random vector $\xi := \sqrt{\rho}\,\mathcal{R}S$. The distribution of ξ is symmetric and so the lower EDC coincides with the upper EDC of ξ. Moreover, the marginal distribution functions of ξ are identical and so the EDC can be calculated by

$$\varepsilon = \lim_{r\to\infty} \frac{\mathbb{P}(\xi > r1)}{1 - \mathbb{P}(\xi \leq r1)} = \lim_{r\to\infty} \frac{\mathbb{P}(\mathcal{R}Y > r)}{\mathbb{P}(\mathcal{R}Z > r)}.$$

By applying the Law of Total Probability we obtain

$$\varepsilon = \lim_{r\to\infty} \frac{\int_0^\infty \mathbb{P}(\mathcal{R} > r/y)\, dF_Y(y)}{\int_0^\infty \mathbb{P}(\mathcal{R} > r/z)\, dF_Z(z)} = \lim_{r\to\infty} \frac{\int_0^\infty \mathbb{P}(\mathcal{R} > y^{-1}r)/\mathbb{P}(\mathcal{R} > r)\, dF_Y(y)}{\int_0^\infty \mathbb{P}(\mathcal{R} > z^{-1}r)/\mathbb{P}(\mathcal{R} > r)\, dF_Z(z)}.$$

Since $\mathcal{R}$ is regularly varying with tail index α, we have that

$$\frac{\mathbb{P}(\mathcal{R} > y^{-1}r)}{\mathbb{P}(\mathcal{R} > r)} \longrightarrow y^\alpha \quad \text{and} \quad \frac{\mathbb{P}(\mathcal{R} > z^{-1}r)}{\mathbb{P}(\mathcal{R} > r)} \longrightarrow z^\alpha, \qquad r \longrightarrow \infty.$$

The convergence is uniform in $(0, a]$ for all $a > 0$ (Embrechts et al., 1997). Further, the distribution of Z and thus also of $Y \leq Z$ has a finite right endpoint. Hence, we can choose a sufficiently large number a and apply the Dominated Convergence Theorem, which implies that

$$\lim_{r\to\infty} \int_0^\infty \frac{\mathbb{P}(\mathcal{R} > y^{-1}r)}{\mathbb{P}(\mathcal{R} > r)}\, dF_Y(y) = \int_0^\infty y^\alpha \, dF_Y(y) < \infty$$

as well as

$$\lim_{r\to\infty} \int_0^\infty \frac{\mathbb{P}(\mathcal{R} > z^{-1}r)}{\mathbb{P}(\mathcal{R} > r)}\, dF_Z(z) = \int_0^\infty z^\alpha \, dF_Z(z) < \infty.$$

This leads us to the desired result. □

The result of Theorem 1 can be expressed, equivalently, by

$$\varepsilon = \frac{\mathbf{E}((\min \sqrt{\rho}\,S \vee 0)^\alpha)}{\mathbf{E}((\max \sqrt{\rho}\,S \vee 0)^\alpha)}, \tag{1}$$

where $a \vee b$ denotes the maximum of $a, b \in \mathbb{R}$. This expression clearly reveals that the EDC of a regularly varying elliptically distributed random vector depends only on ρ and α. In particular, the given formula is comfortable if we want to approximate ε by numerical simulation.

Table 1 contains the EDC of a 2-dimensional regularly varying elliptically distributed random vector for different values of $\rho_{12} = \rho_{21}$ and α. The EDC equals 1 if $\rho_{12} = 1$ or $\alpha = 0$, whereas it equals 0 if $\rho_{12} = -1$ (but not $\alpha = 0$) or $\alpha = \infty$ (but not $\rho_{12} = 1$), where "$\rho_{12} = -1$," "$\rho_{12} = 1$," and "$\alpha = \infty$" shall be interpreted as the limiting cases $\rho_{12} \searrow -1$, $\rho_{12} \nearrow 1$, and $\alpha \to \infty$.

Table 1. EDC for different values of $\boldsymbol{\rho_{12}}$ and $\boldsymbol{\alpha}$ in the case of $\boldsymbol{d = 2}$.

	α							
ρ_{12}	0	1	2	3	4	5	10	∞
-1	1	0	0	0	0	0	0	0
-0.75	1	0.0334	0.0099	0.0031	0.0010	0.0003	0	0
-0.5	1	0.0718	0.0297	0.0130	0.0059	0.0027	0.0001	0
-0.25	1	0.1170	0.0590	0.0316	0.0175	0.0099	0.0006	0
0	1	0.1716	0.0999	0.0616	0.0393	0.0255	0.0034	0
0.25	1	0.2404	0.1576	0.1088	0.0775	0.0563	0.0132	0
0.5	1	0.3333	0.2430	0.1852	0.1449	0.1155	0.0427	0
0.75	1	0.4776	0.3883	0.3261	0.2793	0.2424	0.1338	0
1	1	1	1	1	1	1	1	1

3. Empirical investigation

The EDC is an asymptotic risk measure. Usually, such kind of risk measures are not easy to estimate if the sample size is small or if the estimator is nonparametric (Frahm et al., 2005). The trick is to use a semiparametric approach, i.e., to combine parametric and nonparametric elements. Here, we adopt this approach by restricting ourselves to elliptical distributions. Hence, we allow for a large number of well-known multivariate distributions, e.g., the normal, the sub-Gaussian α-stable distribution (Rachev and Mittnik, 2000) as well as the symmetric generalized hyperbolic distributions (Barndorff-Nielsen et al., 1982). Regular variation excludes the normal distribution and any other elliptical distribution with exponentially decaying tails. However, in the light of Figure 1 and Figure 2, this restriction is not binding at all in our context.

A quite popular parametric alternative to the multivariate normal distribution is the multivariate t-distribution, which is characterized by the density

$$f(x) = \frac{\Gamma\left(\frac{d+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \cdot \sqrt{\frac{\det\left(\Sigma^{-1}\right)}{(\nu\pi)^d}} \cdot \left(1 + \frac{(x-\mu)'\,\Sigma^{-1}\,(x-\mu)}{\nu}\right)^{-\frac{d+\nu}{2}},$$

where $\nu > 0$ is the number of degrees of freedom and the dispersion matrix Σ is assumed to be positive definite. If the number ν tends to infinity, the multivariate t-distribution approaches the multivariate normal distribution. Further, for $\nu = 1$ we obtain the Cauchy distribution. The class of symmetric generalized hyperbolic distributions contains the multivariate t- and thus the Cauchy distribution as special cases. The expectation of a t-distributed random vector X corresponds to μ if $\nu > 1$, but in the case of $0 < \nu \leq 1$ it does not exist. Further, in the case of $\nu > 2$, the covariance matrix of X equals $\frac{\nu}{\nu-2}\,\Sigma$. It is well-known that the multivariate t-distribution represents a regularly varying distribution with tail index ν. Hence, in order to estimate the location vector, the correlation matrix, and the tail index of log-returns in a parametric way, we can fit a multivariate t-distribution by maximum likelihood (ML).

The EDC can be estimated by using the plug-in approach. For this purpose, we have to choose some appropriate estimators for ρ and α in order to substitute the true parameters with the corresponding estimates. According to Theorem 1, the EDC is a function of ρ and α and with equation 1 it is quite simple to compute the estimate of ε given the estimates of ρ and α.

Let $X_{\cdot j} = \mu + \Lambda \mathcal{R}_j S_j$ be the (7-dimensional) vector of log-returns at Day $j = 1, 2, \ldots, n$. Hence, μ and Σ are constant over time and I assume that the generating variates $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_n$ are identically distributed. The components of the d-dimensional stochastic process $\{X_{\cdot n}\}$ need not be serially independent. It suffices to assume that $\{X_{\cdot n}\}$ is (strictly) stationary and ergodic.

3.1. Estimating the correlation matrix

In order to estimate ρ, we may start with Tyler's M-estimator for Σ (Tyler, 1987a, b), viz.

$$\widehat{\Sigma} = \frac{d}{n} \sum_{j=1}^{n} \frac{(X_j - \hat{\mu})(X_j - \hat{\mu})'}{(X_j - \hat{\mu})'\widehat{\Sigma}^{-1}(X_j - \hat{\mu})}\,, \tag{2}$$

where $\hat{\mu}$ is the estimator for μ that is associated with $\widehat{\Sigma}$ in a natural way (Hettmansperger and Randles, 2002; Tyler, 1987a). This estimator proves to be favorable whenever the data exhibit heavy tails (Frahm, 2004; Frahm and Jeakel, 2010, 2015). Tyler's M-estimator is the most robust estimator for Σ if the distribution of X is elliptical (Tyler, 1987a). If the location vector μ is known, the distribution of $\widehat{\Sigma}$ is not affected at all by the generating variate $\mathcal{R}$. For more details on that topic see the aforementioned references as well as Adrover (1998), Dümbgen and Tyler (2005), Kent and Tyler (1988, 1991), Maronna and Yohai (1990), and Tyler (1983, 1987b).

The estimate of ρ based on Tyler's M-estimator for Σ is contained in Table 2. The given results confirm a well-known phenomenon of empirical finance: Returns on stocks and stock indices are, in general, highly correlated with each other in our global economy. Thus, it is interesting to observe that the correlation between Japan and each other G–7 country is relatively low. In fact, Japan is almost uncorrelated with USA. By contrast, the correlation between France and Germany is almost perfect.

Table 2. Estimate of $\boldsymbol{\rho}$ based on Tyler's M-estimator for scatter.

	Canada	France	Germany	Italy	Japan	UK	USA
Canada	1	0.5659	0.5480	0.5159	0.1644	0.5745	0.6155
France	0.5659	1	0.9205	0.8729	0.1871	0.8317	0.5014
Germany	0.5480	0.9205	1	0.8276	0.1833	0.7921	0.5152
Italy	0.5159	0.8729	0.8276	1	0.1417	0.7475	0.4492
Japan	0.1644	0.1871	0.1833	0.1417	1	0.2011	0.0288
UK	0.5745	0.8317	0.7921	0.7475	0.2011	1	0.4707
USA	0.6155	0.5014	0.5152	0.4492	0.0288	0.4707	1

Alternatively, ρ can be estimated in a parametric way by fitting the multivariate t-distribution to the data. This is done by using the algorithm proposed by Aeschliman et al. (2010), which proves to be very fast and reliable. The result is given in Table 3. We can see that the estimates of ρ are quite similar to those given by Table 2.

Table 3. Estimate of $\boldsymbol{\rho}$ based on the multivariate $\boldsymbol{t}$-distribution.

	Canada	France	Germany	Italy	Japan	UK	USA
Canada	1	0.5465	0.5374	0.4994	0.1562	0.5546	0.6362
France	0.5465	1	0.8922	0.8647	0.1815	0.8234	0.4760
Germany	0.5374	0.8922	1	0.8043	0.1673	0.7652	0.5114
Italy	0.4994	0.8647	0.8043	1	0.1389	0.7369	0.4291
Japan	0.1562	0.1815	0.1673	0.1389	1	0.1928	0.0205
UK	0.5546	0.8234	0.7652	0.7369	0.1928	1	0.4532
USA	0.6362	0.4760	0.5114	0.4291	0.0205	0.4532	1

3.2. Estimating the tail index

Extreme value theory provides many possibilities in order to estimate the tail index of a regularly varying random variable (Embrechts et al., 1997). However, the aforementioned authors clearly advocate the peaks-over-threshold (POT) method (Embrechts et al., 1997). First of all, we have to estimate the realizations of $\mathcal{R}$, which represents a latent variable. The dispersion matrix Σ can be identified only up to some scaling constant $\kappa^2 > 0$ because $X = \mu + \Lambda\mathcal{R}S = \mu + (\kappa\Lambda)(\mathcal{R}/\kappa)S$ for all $\kappa > 0$. Tyler's M-estimator suffers from the same identification problem, since Equation 2 remains valid if we substitute the estimate $\widehat{\Sigma}$ with $\kappa^2\widehat{\Sigma}$ for any $\kappa > 0$.

However, we are not affected by the identification problem. Note that

$$(X-\mu)'(\kappa^2\Sigma)^{-1}(X-\mu) = (\mathcal{R}/\kappa)^2,$$

but the tail index of $\mathcal{R}/\kappa$ does not depend on κ at all. For this reason, we can choose any positive constant κ or, equivalently, any appropriate shape matrix Σ (Frahm, 2009; Paindaveine, 2008). Suppose that $\mathbf{E}(\mathcal{R}^2) < \infty$. We will see later on that this assumption is not too farfetched. In this case, we can assume without loss of generality that Σ is such that $\mathbf{E}(\mathcal{R}^2) = d$, which guarantees that $\mathbf{Var}(X) = \Sigma$. Now, the realization r_j of the generating variate at Day $j = 1, 2, \ldots, n$, i.e., $\mathcal{R}_j$, can be estimated by

$$\hat{r}_j = \sqrt{(x_j - \hat{\mu})'\widehat{\Sigma}^{-1}(x_j - \hat{\mu})}\,, \tag{3}$$

where x_j is the realization of X_j and $\widehat{\Sigma}$ is such that $\frac{1}{n}\sum_{j=1}^{n}\hat{r}_j^2 = 7$. Figure 6 contains the kernel density of $\hat{r}_1^2, \hat{r}_2^2, \ldots, \hat{r}_n^2$ and the χ^2_d-density with $d = 7$ degrees of freedom. Once again, we can see that the normal-distribution hypothesis ($\mathcal{R}^2 = \chi^2_d$) is clearly violated for real data.

The mean-excess plot (Embrechts et al., 1997) based on $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n$ is given by Figure 7, which clearly reveals that $\mathcal{R}$ has a power tail with positive tail index. We may choose $\tau = 4$ as a critical threshold and calculate the excess $\hat{w}_j := \hat{r}_j - 4$ for all $\hat{r}_j > 4$ (Embrechts et al., 1997). The POT-estimator for the tail index α represents an ML-estimator that is based on the assumption that the

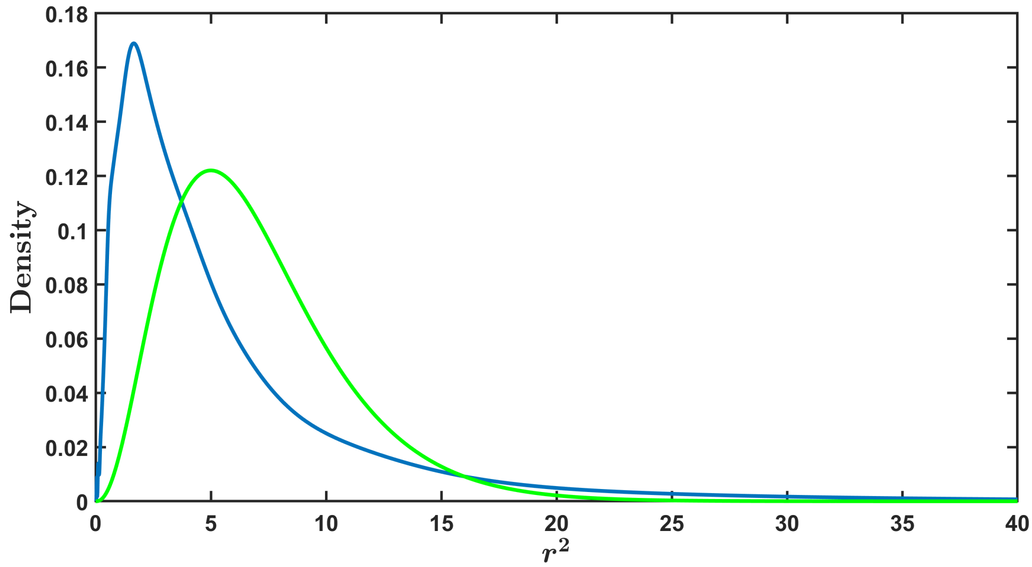

Figure 6. Kernel density of $\hat{r}_1^2, \hat{r}_2^2, \ldots, \hat{r}_n^2$ (blue) vs. the χ_d^2-density with $d = 7$ (green).

excesses follow a generalized Pareto distribution. More precisely, the density of the excess $w \geq 0$ is assumed to be

$$f(w) = \frac{1}{\beta}\left(1 + \frac{w}{\alpha\beta}\right)^{-\alpha-1},$$

where $\alpha > 0$ is the tail index and $\beta > 0$ represents a scale parameter.

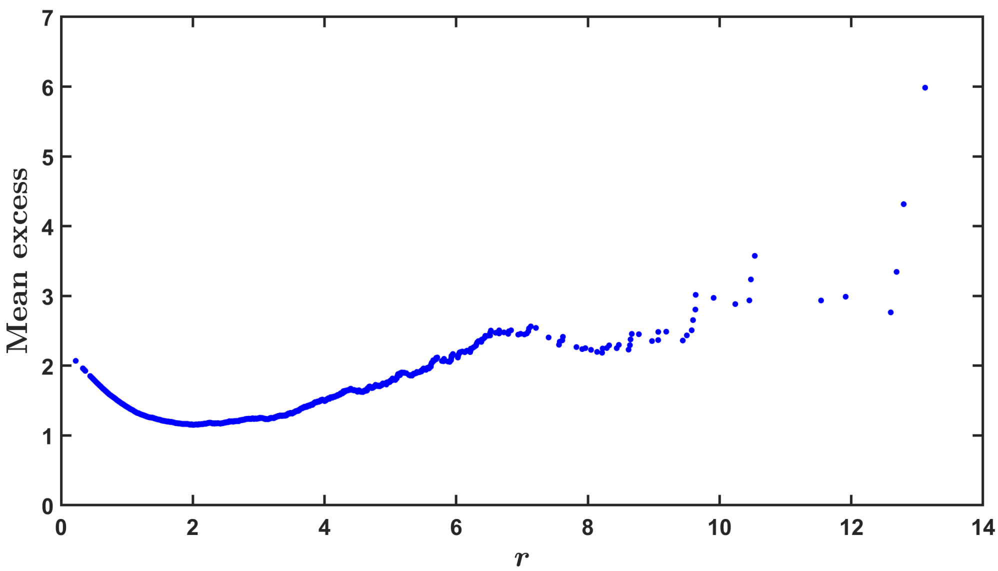

Figure 7. Empirical mean-excess function of $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n$.

After applying the ML-estimator to $\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_n$ we obtain the result $\hat{\alpha} = 4.1705$. The corresponding standard error is 1.1004 under the simplifying assumption that the observations are serially independent. Hence, the estimation risk is very large, which is a typical phenomenon when applying extreme value theory to financial data. The one-sided 95%-confidence interval for α is $[2.3605, \infty)$, whereas the two-sided 95%-confidence interval corresponds to $[2.0137, 6.3273]$. Thus, we may at least expect that $\alpha > 2$, i.e., that the second moment of $\mathcal{R}$ is finite, but not much more.

Alternatively, we can use the parametric estimator for ν based on the assumption that the data are multivariate t-distributed. Recall that ν is the tail index of the multivariate t-distribution. This leads to $\hat{\nu} = 2.9013$ with standard error 0.0577. The standard error of $\hat{\nu}$ is essentially lower than that of $\hat{\alpha}$, but we should keep in mind that the assumption that daily log-returns are multivariate t-distributed can be wrong. To check this, we can calculate the realizations of $\mathcal{R}$, i.e., $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n$, by using Eq. 3. Now, $\hat{\mu}$ and $\widehat{\Sigma}$ correspond to the ML-estimates based on the multivariate t-distribution. In this case, we do not have any identification problem concerning the dispersion matrix Σ and so we need not normalize the realizations of $\mathcal{R}$.

If the log-returns follow a multivariate t-distribution, then $\mathcal{R}^2$ equals $dF_{d,\nu}$, where $F_{d,\nu}$ is an F-distributed random variable with d numerator degrees of freedom and ν denominator degrees of freedom (Frahm, 2004). Hence, in order to get an impression of the goodness of fit, the reader can compare the kernel density of $\hat{r}_1^2, \hat{r}_2^2, \ldots, \hat{r}_n^2$ with the density of $dF_{d,\nu}$ in Figure 8. As we can see, the fit is good in the tail, but it is still unsatisfactory in the center. However, from the viewpoint of a risk manager, the multivariate t-distribution is clearly preferable to the normal distribution (see Figure 6), which seriously underestimates the heaviness of the tail. Nonetheless, when comparing $\hat{\alpha}$ with $\hat{\nu}$ we must be aware of the typical bias-variance trade-off between parametric and nonparametric methods.

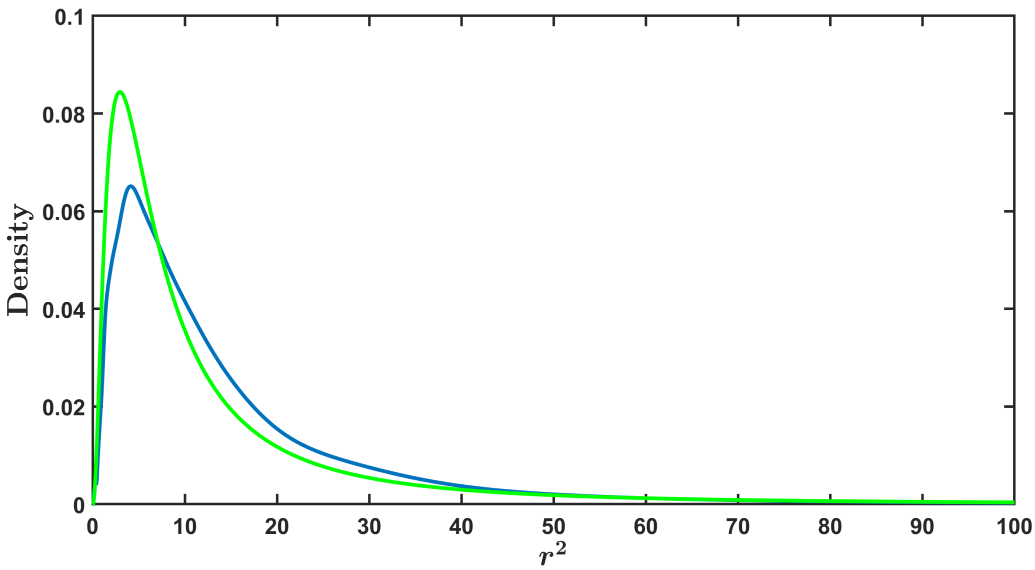

Figure 8. Kernel density of $\hat{r}_1^2, \hat{r}_2^2, \ldots, \hat{r}_n^2$ (blue) vs. the density of $dF_{d,\nu}$ with $d = 7$ and $\nu = 2.9013$ (green).

3.3. Ruin probabilities

In the following, let

- $\pi := \mathbb{P}(\min U \leq p)$ be the probability that at least one country has a p-shortfall and

- $\psi := \mathbb{P}(\max U \leq p)$ be the probability that all countries have a p-shortfall.

The latter is referred to as a one-day ruin probability.

If p is sufficiently small, we have that

$$\psi = \frac{\mathbb{P}(\max U \leq p)}{\mathbb{P}(\min U \leq p)} \cdot \mathbb{P}(\min U \leq p) \approx \varepsilon\pi\,.$$

This simple approximation can be used in order to estimate the ruin probability ψ by $\hat{\varepsilon}\hat{\pi}$, where $\hat{\varepsilon}$ is the plug-in estimator for ε and $\hat{\pi}$ is the empirical estimator for π. The basic idea is to use an empirical estimator whenever the number of observations is large enough, but to apply a semiparametric approach if the number of observations is small or even zero. Estimating π is easy because the number of trading days on which *at least* one country had a shortfall is relatively large. Even for the smallest shortfall probability $p = 0.005$ we can observe 77 out of 4804 days that satisfy this condition. By contrast, the number of trading days on which *all* countries had a shortfall is very small. There was only *one* day on which all G–7 countries had a 0.005-shortfall, i.e., November 6, 2008. In many practical applications this number can even be zero. Of course, this makes nonparametric estimation of ψ impossible. For this reason, we apply a semiparametric approach in order to estimate ε.

Table 4 contains the empirical 20-day, 100-day, and 200-day quantiles for the G–7 countries and the (estimated) probabilities that at least one country suffers from an associated shortfall. Canada, Japan, and UK seem to have quite similar quantiles and the same holds true for France, Germany, and Italy. The given numbers are just estimates and so it is clear that they suffer from estimation risk. I cannot see any economic argument that explains why these countries should have similar *theoretical* quantiles.

Table 4. Empirical 20-, 100-, and 200-day quantiles for the G–7 countries.

p	Canada	France	Germany	Italy	Japan	UK	USA	π
0.05	-0.0217	-0.0245	-0.0255	-0.0257	-0.0222	-0.0210	-0.0187	0.1559
0.01	-0.0394	-0.0456	-0.0468	-0.0507	-0.0379	-0.0369	-0.0346	0.0339
0.005	-0.0496	-0.0544	-0.0586	-0.0603	-0.0465	-0.0462	-0.0435	0.0160

During m trading days we can expect

$$\mathbf{E}\left(\sum_{j=1}^{m} \mathbf{1}_{\max U_{\cdot j} \leq p}\right) = \sum_{j=1}^{m} \mathbb{P}(\max U_{\cdot j} \leq p) = m\psi$$

ruins, where $U_{\cdot j} = (U_{1j}, U_{2j}, \dots, U_{dj})$ with $U_{ij} := F_i(X_{ij})$ for $i = 1, 2, \dots, d$ and $j = 1, 2, \dots, n$. This formula holds irrespective of whether the shortfalls are serially dependent or independent. That is, on average, a ruin occurs after ψ^{-1} trading days, i.e., $\psi^{-1}/250$ years, and so this is referred as to the expected ruin time.

Let ψ_m be the m-day ruin probability, i.e., the probability of a financial collapse during m trading days. If we make the simplifying assumption that the ruins are serially independent, we obtain

$$\psi_m = 1 - (1 - \psi)^m \approx 1 - (1 - \varepsilon\pi)^m \tag{4}$$

for all $m \in \{1, 2, \dots\}$. However, as already mentioned in Section 1, in real life we can observe shortfall clusters and so the serial-independence assumption is violated. We can expect that the probability of subsequent drawdowns increases in turbulent times and decreases when the financial market is calm. In the finance literature, this phenomenon is often described by so-called Hawkes processes, i.e., self-exciting point processes (Laub et al., 2015). Nonetheless, for the sake of simplicity, here I assume that concomitant shortfalls are serially independent.

Risk managers generally distinguish between the short-term and the long-term approach. The short-term approach is dynamic and refers to the conditional distribution of asset returns, whereas the long-term approach is static and refers to their unconditional distribution. Whenever we apply empirical methods of risk management, we should make clear from the outset whether our approach is dynamic or static. The approach advocated in this work is clearly static because it takes the unconditional, i.e., stationary, distribution of asset returns into account. This means that it is not appropriate if we want to quantify the systemic risk during the *forthcoming* five years.[§] In order to accomplish the ambitious task of a dynamic forecast, we would have to take the current state of the economy into consideration and use suitable parametric models, i.e., models that reflect the dynamic properties of daily log-returns in a reasonable way. This goes far beyond the present work. Nonetheless, I am convinced that the methods proposed here can readily be used also in a conditional rather than unconditional framework.

The POT-estimate for the tail index α is roughly 4. It is worth emphasizing that this result does not change substantially if we use another estimator for the tail index, e.g., the Hill estimator or the Pickands estimator (Embrechts et al., 1997). For this reason, we can conduct a scenario analysis with $\alpha = 4$ representing the normal case. By contrast, due to the confidence interval for α reported in Section 3.2, the tail index $\alpha = 2$ represents the worst case, whereas $\alpha = 6$ is the best case. Table 5 contains the results of our analysis. Here, I recapitulate the given parameters:

- p is the shortfall probability,
- ε is the EDC,
- π is the probability that at least one country has a p-shortfall,
- ψ is the probability that all countries have a p-shortfall,
- $\psi^{-1}/250$ is the expected ruin time (in years),
- ψ_{250} is the 1-year ruin probability,
- $\psi_{5\cdot 250}$ is the 5-year ruin probability, and
- $\psi_{10\cdot 250}$ is the 10-year ruin probability.

Table 6 contains the parameter estimates based on the multivariate t-distribution. If the daily log-returns were multivariate t-distributed, we should trust these results more than those in Table 5, but as we have already seen, the multivariate t-distribution is not beyond all doubt.

Figure 9 illustrates how ruin probabilities, based on the shortfall probability $p = 0.01$, depend on the tail index α. We can see that the ruin probabilities are quite sensitive to the tail index. Obtaining a valid semiparametric estimate of α is a challenge because we have to cope with the general bias-variance trade-off that is well-known in extreme value theory. However, for a shortfall probability of $p = 0.05$, the given results clearly indicate that in the normal case ($a = 4$) we will observe a collapse of the entire financial market each 5 years. In the best case ($\alpha = 6$) the expected ruin time is much longer and in the worst case ($\alpha = 2$) it is much shorter. It is very unlikely that the tail index is below 2 because then the number of collapses would have been much larger during the last decades. Hence, the tails of sub-Gaussian α-stable distributions appear to be too heavy, which confirms a similar result concerning

[§]By the way, the same issue arises whenever we try to estimate the value at risk.

Table 5. Analytical results based on the POT-estimator.

$\hat{\alpha} = 4.1705$ (POT-estimate), standard error: 1.1004							
p	ε	π	ψ	$\psi^{-1}/250$	ψ_{250}	$\psi_{5\cdot250}$	$\psi_{10\cdot250}$
0.05	0.0041	0.1559	0.0006	6.1852	0.1493	0.5545	0.8016
0.01	0.0041	0.0339	0.0001	28.4698	0.0345	0.1611	0.2962
0.005	0.0041	0.0160	0.0001	59.8322	0.0166	0.0802	0.1539
$\alpha = 2$ (worst case)							
p	ε	π	ψ	$\psi^{-1}/250$	ψ_{250}	$\psi_{5\cdot250}$	$\psi_{10\cdot250}$
0.05	0.0192	0.1559	0.0030	1.3345	0.5274	0.9764	0.9994
0.01	0.0192	0.0339	0.0007	6.1322	0.1502	0.5569	0.8036
0.005	0.0192	0.0160	0.0003	12.9811	0.0739	0.3189	0.5361
$\alpha = 4$ (normal case)							
p	ε	π	ψ	$\psi^{-1}/250$	ψ_{250}	$\psi_{5\cdot250}$	$\psi_{10\cdot250}$
0.05	0.0046	0.1559	0.0007	5.5232	0.1642	0.5921	0.8336
0.01	0.0046	0.0339	0.0002	25.3796	0.0382	0.1771	0.3229
0.005	0.0046	0.0160	0.0001	53.7256	0.0182	0.0879	0.1681
$\alpha = 6$ (best case)							
p	ε	π	ψ	$\psi^{-1}/250$	ψ_{250}	$\psi_{5\cdot250}$	$\psi_{10\cdot250}$
0.05	0.0014	0.1559	0.0002	18.7455	0.0531	0.2388	0.4206
0.01	0.0014	0.0339	< 0.0001	86.1372	0.0118	0.0576	0.1119
0.005	0.0014	0.0160	< 0.0001	182.3423	0.0056	0.0276	0.0545

the TDC reported by Frahm et al. (2003). That is, we can expect that the log-returns have a (finite) covariance matrix. The probability of a ruin during some relatively short period of time, e.g., 5 or 10 years, turns out to be high from a risk-manager's perspective and we cannot exclude the possibility that the tail index, α, increases during the coming decades. However, in this work, I assume that $\{X_{\cdot n}\}$ is stationary and thus α is constant. Testing for structural breaks concerning the tail index would require us to analyze much longer time series. This could be the subject of future research.

Table 6. Analytical results based on the multivariate $\boldsymbol{t}$-distribution.

$\hat{\nu} = 2.9013$ (ML-estimate), standard error: 0.0577							
p	ε	π	ψ	$\psi^{-1}/250$	ψ_{250}	$\psi_{5\cdot250}$	$\psi_{10\cdot250}$
0.05	0.0098	0.1559	0.0015	2.6098	0.3185	0.8530	0.9784
0.01	0.0098	0.0339	0.0003	12.0678	0.0795	0.3393	0.5634
0.005	0.0098	0.0160	0.0002	25.4897	0.0385	0.1781	0.3245

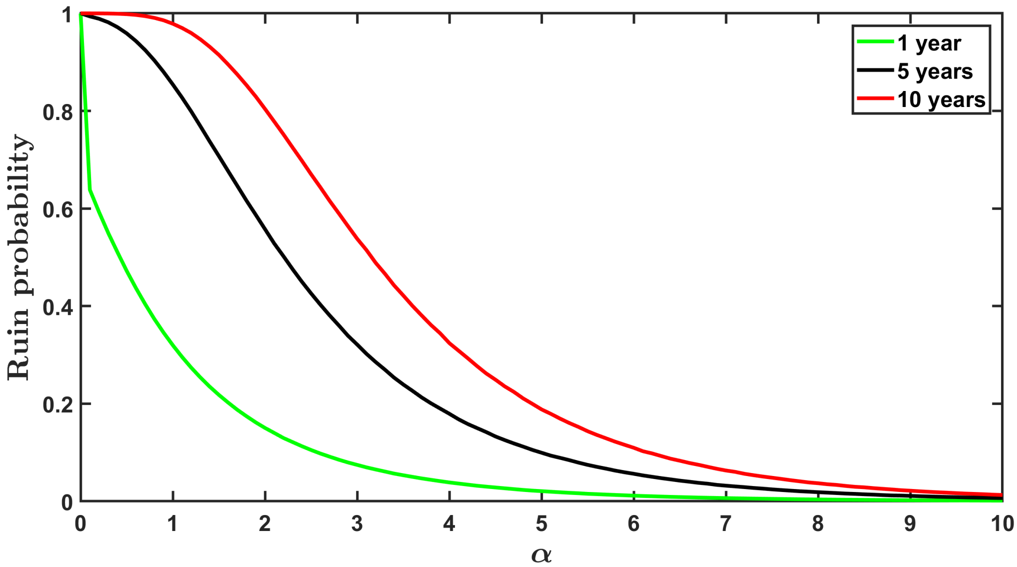

Figure 9. Ruin probabilities based on $\boldsymbol{p} = \mathbf{0.01}$ for 1, 5, and 10 years.

4. Conclusions

Daily asset returns are heavy tailed and extremal dependent, which can be described by the assumption that they are regularly varying and elliptically distributed. Copula theory proves suitable for analyzing extremal dependence in a general framework, whereas extreme value theory provides the necessary tools in order to quantify the dependence structure of extreme asset returns that stem from a regularly varying elliptical distribution. The EDC is a natural measure of systemic risk and the EDC of regularly varying elliptically distributed asset returns depends only on their correlation matrix and the tail index. Extreme value theory allows us to estimate the EDC in a semiparametric way. This point is essential because, by their very definition, extreme values do not appear often in real life and so it is virtually impossible to apply a purely nonparametric estimator in order to estimate the EDC.

The presented theory has been applied in order to analyze the risk that the financial market collapses. The indicated ruin probabilities are high, but we must notice that the results are quite sensitive to the tail index. In order to take different tail indices into account, we conducted a scenario analysis. In the worst case, the probability that the financial market collapses during 10 years exceeds 50%. Hence, at least from a risk-manager's perspective, we should remain cautious and must not neglect the risk that international diversification dramatically fails from time to time. Nonetheless, at least we can reject the hypothesis that daily asset returns have no finite second moments, which precludes the sub-Gaussian α-stable distribution. This confirms a similar result obtained by Frahm et al. (2003) regarding the TDC.

5. Conflict of interest

The authors declare no conflict of interest.

References

Adrover J (1998) Minimax bias-robust estimation of the dispersion matrix of a multivariate distribution. *Ann Stat* 26: 2301–2320.

Aeschliman C, Park J, Kak A (2010) A novel parameter estimation algorithm for the multivariate t-distribution and its application to computer vision, in K. Daniilidis, P. Maragos, N. Paragios (editors), *European Conference on Computer Vision 2010*, Springer, 594–607.

Artzner P, Delbaen F, Eber JM, et al. (1999) Coherent measures of risk. *Math Financ* 9: 203–228.

Barndorff-Nielsen O, Kent J, Sørensen M (1982) Normal variance-mean mixtures and z distributions. *Int Stat Rev* 50: 145–159.

Bingham N, Kiesel R (2002) Semi-parametric modelling in finance: theoretical foundation. *Quant Financ* 2: 241–250.

Bouchaud JP, Cont R, Potters M (1997) Scaling in stock market data: stable laws and beyond, in B. Dubrulle, F. Graner, D. Sornette (editors), *Scale Invariance and Beyond*, Proceedings of the CNRS Workshop on Scale Invariance, Les Houches, March 1997, EDP-Springer.

Breymann W, Dias A, Embrechts P (2003) Dependence structures for multivariate high-frequency data in finance. *Quant Financ* 3: 1–14.

Cambanis S, Huang S, Simons G (1981) On the theory of elliptically contoured distributions. *J Multivariate Anal* 11: 368–385.

Cont R (2001) Empirical properties of asset returns: stylized facts and statistical issues. *Quant Financ* 1: 223–236.

De Luca G, Rivieccio G (2012) Multivariate tail dependence coefficients for Archimedean copulae, in A. Di Ciaccio, M. Coli, J. Ibanez(editors), Advanced Statistical Methods for the Analysis of Large Data-Sets, Studies in Theoretical and Applied Statistics, Springer, 287–296.

Ding Z, Granger C, Engle R (1993) A long memory property of stock market returns and a new model. *J Empir Financ* 1: 83–106.

Dobrić J, Frahm G, Schmid F (2013) Dependence of stock returns in bull and bear markets. *Dependence Modeling* 1: 94–110.

Dümbgen L, Tyler D (2005) On the breakdown properties of some multivariate M-functionals. *Scand J Stat* 32: 247–264.

Eberlein E, Keller U (1995) Hyperbolic distributions in finance. *Bernoulli* 1: 281–299.

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

Engle R (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* 50: 987–1007.

Fama E (1965) The behavior of stock market prices. *J Bus* 38: 34–105.

Fang K, Kotz S, Ng K (1990) Symmetric Multivariate and Related Distributions, Chapman & Hall.

Ferreira H, Ferreira M (2012) On extremal dependence of block vectors. *Kybernetika* 48: 988–1006.

Frahm G (2004) Generalized Elliptical Distributions: Theory and Applications, Ph.D. thesis, University of Cologne.

Frahm G (2006) On the extremal dependence coefficient of multivariate distributions. *Stat Probabil Lett* 76: 1470–1481.

Frahm G (2009) Asymptotic distributions of robust shape matrices and scales. *J Multivariate Anal* 100: 1329–1337.

Frahm G, Jaekel U (2010) A generalization of Tyler's M-estimators to the case of incomplete data. *Comput Stat Data An* 54: 374–393.

Frahm G, Jaekel U (2015) Tyler's M-estimator in high-dimensional financial-data analysis, in K. Nordhausen, S. Taskinen (editors), *Modern Nonparametric, Robust and Multivariate Methods*, Chapter 17, Springer, 289–305.

Frahm G, Junker M, Schmidt R (2005) Estimating the tail-dependence coefficient: properties and pitfalls. *Insur Math Econ* 37: 80–100.

Frahm G, Junker M, Szimayer A (2003) Elliptical copulas: applicability and limitations. *Stat Probabil Lett* 63: 275–286.

Hettmansperger T, Randles R (2002) A practical affine equivariant multivariate median *Biometrika* 89: 851–860.

Hult H, Lindskog F (2002) Multivariate extremes, aggregation and dependence in elliptical distributions. *Adv Appl Probab* 34: 587–608.

Joe H (1997) Multivariate Models and Dependence Concepts, Chapman & Hall.

Jorion P (1985) International portfolio diversification with estimation risk. *J Bus* 58: 259–278.

Junker M, May A (2005) Measurement of aggregate risk with copulas. *Economet J* 8: 428–454.

Kelker D (1970) Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhya A* 32: 419–430.

Kent J, Tyler D (1988) Maximum likelihood estimation for the wrapped Cauchy distribution. *J Appl Stat* 15: 247–254.

Kent J, Tyler D (1991) Redescending M-estimates of multivariate location and scatter. *Ann Stat* 19: 2102–2119.

Laub P, Taimre T, Pollett P (2015) Hawkes processes, Technical report, Cornell University Library, arXiv:1507.02822.

Mandelbrot B (1963) The variation of certain speculative prices. *J Bus* 36: 394–419.

Markowitz H (1952) Portfolio selection. *J Financ* 7: 77–91.

Maronna R, Yohai V (1990) The maximum bias of robust covariances. *Commun Stat-Theor M* 19: 3925–3933.

McNeil A, Frey R, Embrechts P (2005) Quantitative Risk Management, Princeton University Press.

Mikosch T (2003) Modeling dependence and tails of financial time series, in B. Finkenstaedt, H. Rootzén (editors), Extreme Values in Finance, Telecommunications, and the Environment, Chapman & Hall.

Nelsen R (2006) An Introduction to Copulas, Springer, 2nd edition.

Paindaveine D (2008) A canonical definition of shape. *Stat Probabil Lett* 78: 2240–2247.

Rachev S, Mittnik S (2000) Stable Paretian Models in Finance, Wiley.

Schmidt R (2002) Tail dependence for elliptically contoured distributions. *Math Method Oper Res* 55: 301–327.

Shorrocks A, Davies J, Lluberas R (2017) Global Wealth Databook 2017, Technical report, Credit Suisse Research Institute.

Sklar A (1959) Fonctions de répartition à *n* dimensions et leurs marges, Research report, University of Paris, Institute of Statistics, France.

Tyler D (1983) Robustness and efficiency properties of scatter matrices *Biometrika* 70: 411–420.

Tyler D (1987a) A distribution-free M-estimator of multivariate scatter *Ann Stat* 15: 234–251.

Tyler D (1987b) Statistical analysis for the angular central Gaussian distribution on the sphere *Biometrika* 74: 579–589.

Appendix

QQ–Plots

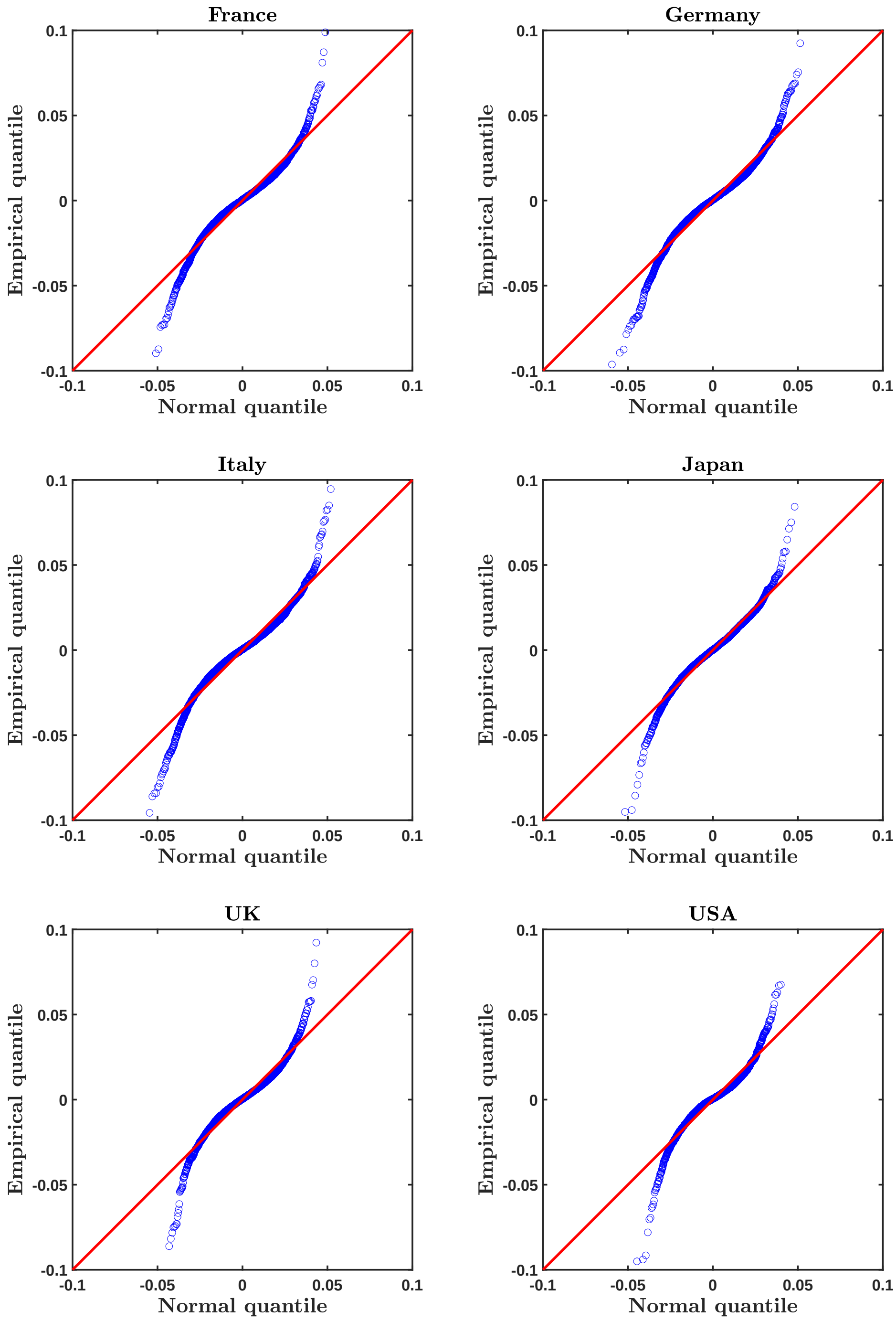

Figure 10. Normal Q–Q plots of daily log-returns from 1999-01-04 to 2018-03-02 on the MSCI country indices except for Canada (see Figure 1).

Scatter Plots

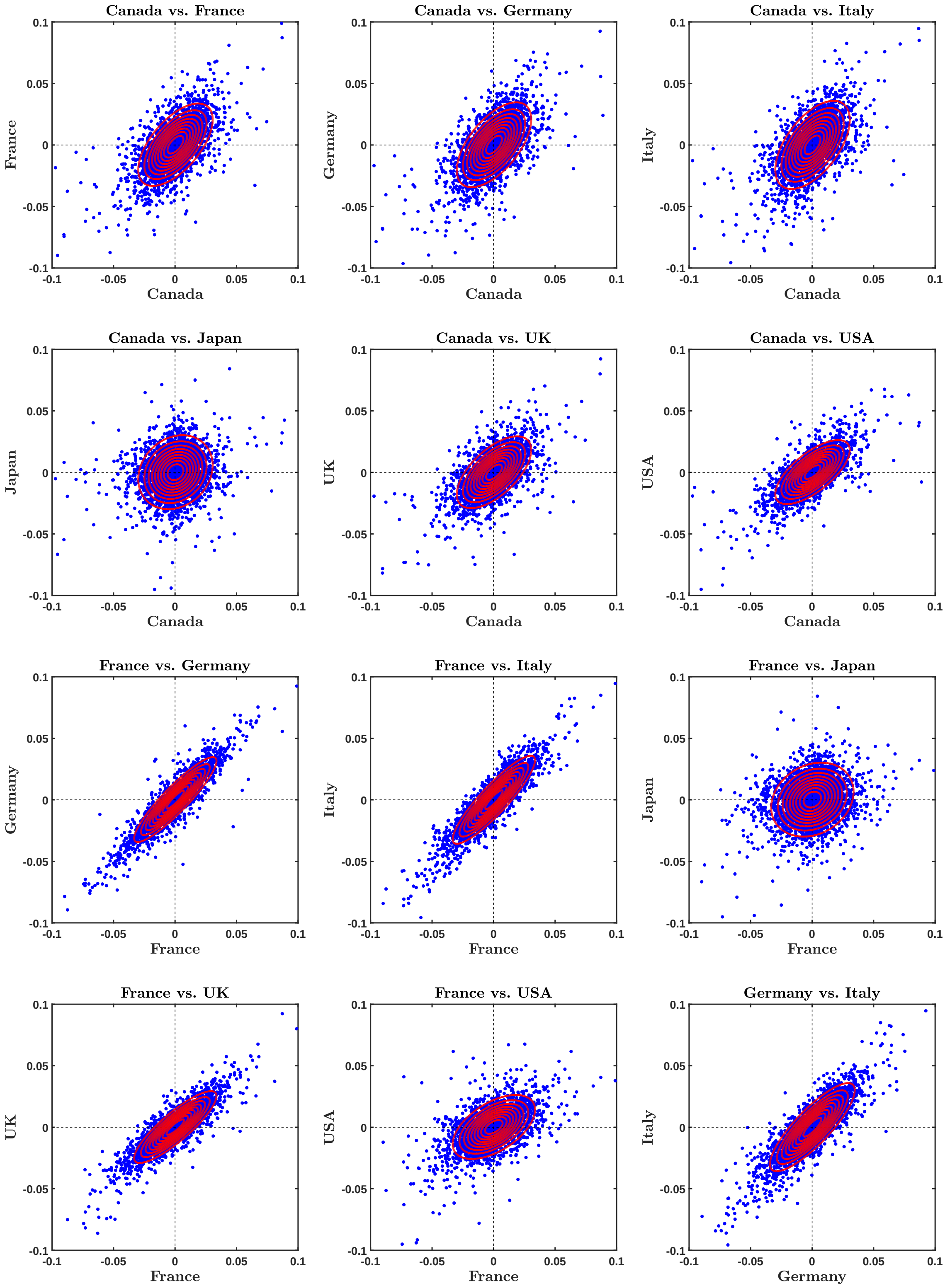

Figure 11 (continued)

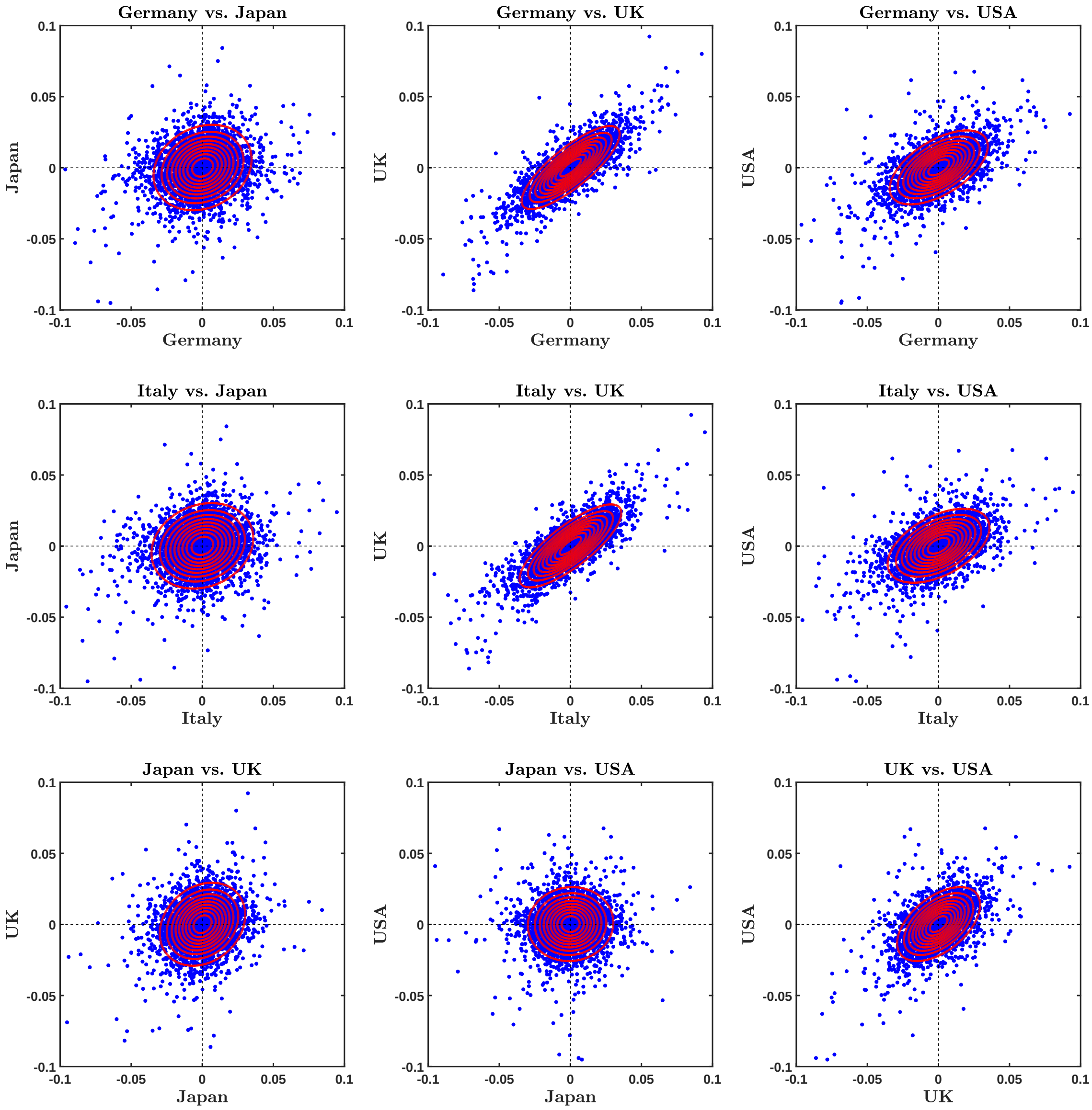

Figure 11. Scatter plots of daily log-returns from 1999-01-04 to 2018-03-02 on the MSCI country indices for all G–7 countries (blue points). The red contours represent the deciles of the bivariate normal distribution that is fitted to the data.

AIMS Press