COLOGNE GRADUATE SCHOOL
in Management, Economics and Social Sciences

Vol 3 | No 6

CGS WORKING PAPER

Modeling Multivariate Extreme Events Using Self-Exciting Point Processes

Oliver Grothe (University of Cologne)
Volodymyr Korniichuk (University of Cologne)
Hans Manner (University of Cologne)

Cologne Graduate School
in Management, Economics
and Social Sciences
Albertus-Magnus-Platz
50923 Köln
www.cgs.uni-koeln.de

University of Cologne

Modeling Multivariate Extreme Events Using Self-Exciting Point Processes

Oliver Grothe[a], Volodymyr Korniichuk[*,b], Hans Manner[a]

[a] *University of Cologne, Department of Economic and Social Statistics*
[b] *University of Cologne, Cologne Graduate School*

Abstract

We propose a new model that can capture the typical features of multivariate extreme events observed in financial time series, namely clustering behavior in magnitudes and arrival times of multivariate extreme events, and time-varying dependence. The model is developed in the framework of the peaks-over-threshold approach in extreme value theory and relies on a Poisson process with self-exciting intensity. We discuss the properties of the model, treat its estimation, deal with testing goodness-of-fit, and develop a simulation algorithm. The model is applied to return data of two stock markets and four major European banks.

Keywords: Time Series, Peaks Over Threshold, Hawkes Processes, Extreme Value Theory.

JEL classification: C32, C51, C58, G15.

1. Introduction

A characteristic property of multivariate financial time series is clustering of univariate and joint extreme returns. While there exist univariate models, which combine extreme value theory and inhomogeneous point processes to cope with the clustering of extremes, models for the multivariate case are still unexplored.

[*] Corresponding author

Email addresses: grothe@statistik.uni-koeln.de (Oliver Grothe), korniichuk@wiso.uni-koeln.de (Volodymyr Korniichuk), manner@statistik.uni-koeln.de (Hans Manner)

June 27, 2012

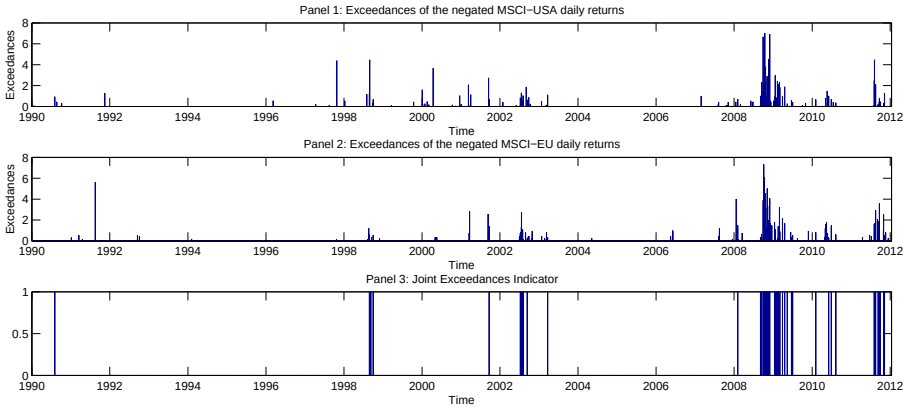

Figure 1: Exceedances of negated MSCI-USA (Panel 1) and MSCI-EU (Panel 2) daily log-returns over the respective 0.977th quantiles. Bar plot indicating times of the joint exceedances (Panel 3).

As an empirical illustration consider Figure 1, which illustrates times and magnitudes of exceedances over a high quantile of negated daily returns of the MSCI-USA and MSCI-EU indices. It is apparent that the exceedances are not randomly distributed, but occur in clusters. Both the times and the magnitudes of exceedances resemble a certain clustering behavior, namely large values tend to be followed by large values and vice versa. Additionally, clusters tend to occur simultaneously in both markets. It is not clear from the figure whether the joint extremes are triggered by one of the markets, i.e., a leading market, or just happen at the same time being caused by a common factor.

In this paper we develop a multivariate approach to model joint exceedances over high thresholds. The model gives insight into the temporal- as well as cross-dependence structure of multivariate extremes. It provides probabilities of joint exceedances conditional on the past of the process, captures clustering behavior in joint exceedances, and accounts for the fact that not only joint but also marginal exceedances may trigger subsequent joint extremes. Furthermore, the model captures asymmetric influences of marginal exceedances so that spill-over and contagion effects in financial market may be analyzed.

The approach is based on a combination of a point process approach to multi-

variate extremes and Hawkes processes (Hawkes, 1971). It captures the times, as well as the marks of both joint and marginal exceedances. We model the behavior of the data explicitly only above a high threshold. Therefore, we do not need to specify time series models for the non-extreme parts of the data. We further specify the dependence structure of joint exceedances only in regions where the results from multivariate extreme value theory (MEVT) are valid. Due to the wide applicability of the limit results of MEVT the proposed model is relatively robust against misspecifications. Furthermore, the MEVT enables to extrapolate exceedance probabilities far beyond the initial threshold, which is needed when one is interested in remote regions of the tail where hardly any data is available.

Modeling multivariate exceedances over high thresholds is a topic of intensive research in extreme value theory. A bivariate analogue of the point over threshold model was presented in Smith, Tawn, and Coles (1997) and a multivariate version in Rootzen and Tajvidi (2006). It was shown that the multivariate Generalized Pareto Distribution (GPD) is the natural distribution for multivariate extreme exceedances, because it bares similar properties as the univariate GPD. Extensive references can be found in de Haan and Ferreira (2006) or Resnick (2007). However, the methods are not directly applicable when the extremes are clustering in time, because the assumption of independent realizations is violated.

An approach for modeling clustered multivariate exceedances over high thresholds, which does not rely on the extreme value theory, is a direct application of multivariate Hawkes process. This has been done, for instance, in Embrechts, Liniger, and Lin (2011), Bowsher (2007), Errais, Giesecke, and Goldberg (2010), or Ait-Sahalia, Cacho-Diaz, and Laeven (2011). In contrast to the models used in these studies ours has several advantages. First, by relying on arguments from extreme value theory it accounts for the specific behavior in the tails of the distribution and allows for extrapolating the exceedance probabilities into the remote tail. This allows for more flexibility concerning the choice of the threshold. Second, our model specification is reasonably parsimonious in terms of the number of model parameters, even when the dimension is large, while maintaining a high degree of flexibility. Finally, we model the intensities of marginal and joint exceedances, while the approaches in the

above mentioned papers only model one of them. In particular, information on both marginal and joint exceedances is allowed to drive the intensity of joint exceedances, therefore capturing possible (asymmetric) spillover effect.

The rest of the paper is structured as follows. The model is derived in Section 2. In Section 3 we describe estimation of the model, along with the goodness-of-fit and simulation procedures. Section 4 presents applications of the model to financial data and Section 5 concludes.

2. Model

In this section we introduce our approach to model the probability of joint extreme events of random vectors $\mathbf{X}_t = (X_{1,t}, \ldots, X_{d,t})$, $t = 1, 2, \ldots, n$, conditioned on the history of the past realizations. At this point, we define an extreme event as the occurrence of a large realization of $X_{i,t}$, $i = 1, \ldots, d$ that lies above an *initial threshold* u_i, e.g., the $95th$ quantile of the marginal distribution. We refer to such extreme events as "exceedances". In our model the exceedances are subject to extreme value theory (EVT), accounting for the specific behavior of the excesses in the tail of the distribution, and to self-exciting jumps, accounting for the clustering and spillover effects in times of occurrence of extreme events observed in real data.

Throughout the text we use the following notation. By $\mathbf{u} = (u_1, \ldots, u_d)$, the *initial threshold vector*, we denote a vector with components relating to sufficiently high quantiles of the marginal distributions of $\mathbf{X}_t$. Modeling the probability of joint exceedances over $\mathbf{x} = (x_1, \ldots, x_d)$ we focus on $\mathbf{x} \geq \mathbf{u}$ (compenentwise). We denote the history of exceedances of $X_{i,1}, \ldots, X_{i,t}$ over u_i by $H_{i,t}$. The history includes both the times $T_{i,1}, \ldots, T_{i,N_{u_i}}$, where N_{u_i} is the number of exceedances with $X_i > u_i$, and the magnitudes $\tilde{X}_{i,1}, \ldots, \tilde{X}_{i,N_{u_i}}$ (the *marks*) of the exceedances. The union of the sets $H_{1,t}, \ldots, H_{d,t}$ is denoted by H_t. With this notation the conditional probability of the joint exceedances in which we are interested can be expressed as

$$P(X_{1,t+1} \geq x_1, \ldots, X_{d,t+1} \geq x_d \mid H_t). \tag{1}$$

Note that being interested in the probability of joint exceedances, the conditioning

set in (1) includes joint and marginal exceedances. Using this extended conditioning set accounts for the fact that both individual and joint extremes influence the arrival of future multivariate extreme events. Before introducing our multivariate model we therefore review the univariate approach to modeling exceedances. In Section 2.2 we then present our model, where we extend the univariate approach to the bivariate case and in Section 2.3 we present a general multidimensional specification.

2.1. Univariate Model

The basic setup to model univariate exceedances is to assume independent and identically distributed (iid) data, i.e., time constant exceedance probabilities, and to use a peaks-over-threshold (POT) model, see, e.g., McNeil, Frey, and Embrechts (2005). The POT model is based on the asymptotic behavior of the threshold exceedances for iid data if these are in the maximum domain of attraction of some extreme value distribution. If the threshold is high enough, then the exceedances occur in time according to a homogeneous Poisson process and the mark sizes are independently and identically distributed according to the generalized Pareto distribution (GPD).

In particular, assume that the marginal distribution functions F_i, with $i = 1, \ldots, d$, of the random vector $\mathbf{X} = (X_1, \ldots, X_d)$ are in the maximum domain of attraction (MDA) of some extreme value distribution G_i with extreme value index ξ_i, denoted as $F_i \in MDA(\xi_i)$. This implies that there exist $a_{i,n} > 0$ and $b_{i,n} \in \mathbb{R}$ such that for $x_{i,n} = a_{i,n}x_i + b_{i,n}$ it holds that

$$\lim_{n\to\infty} nP(X_i \geq x_{i,n}) = \lim_{n\to\infty} n(1 - F_i(x_{i,n})) = -\log G_i(x_i), \tag{2}$$

where G_i is the extreme value distribution, so

$$-\log G_i(x_i) = (1 + \xi_i x_i)^{-1/\xi_i}, \tag{3}$$

for all x_i with $1 + \xi_i x_i > 0$. For $\xi_i = 0$ the right-hand side of (3) is interpreted as e^{-x_i}.

Equation (2) is the standard Poisson limit result, see, e.g., Leadbetter (1991),

which suggests a homogeneous Poisson process with rate $\tau_i(u_i)$ as an appropriate model to describe the occurrence of exceedances of the level u_i by n independent realizations of X_i. In this case the rate is approximated as

$$\tau_i(u_i) = \frac{1}{n}\left(-\log G_i\left(\frac{u_i-\mu_i}{\sigma_i}\right)\right) = \frac{1}{n}\left(1+\xi_i\frac{u_i-\mu_i}{\sigma_i}\right)^{-1/\xi_i}, \tag{4}$$

where μ_i and σ_i represent, respectively, $b_{i,n}$ and $a_{i,n}$ for some fixed n. For a higher threshold $x_i \geq u_i$, the rate of exceedance scales as

$$\tau_i(x_i) = \tau_i(u_i)\left(1+\xi_i\frac{x_i-u_i}{\beta_i}\right)^{-1/\xi_i}, \tag{5}$$

where $\beta_i = \sigma_i + \xi_i(u_i - \mu_i)$. Note that from (2) and (4) the rate $\tau_i(x_i)$ can also be interpreted as

$$\tau_i(x_i) \approx P(X_{i,t} > x_i), \qquad x_i \geq u_i. \tag{6}$$

This interpretation was also encountered in the context of the self-exciting POT model with predictable marks in McNeil, Frey, and Embrechts (2005).

In *self-exciting POT models* this standard setup is extended to allow for temporal dependence in time series such as clustering effects of extreme events. One distinguishes the cases where the probability of exceedances and their distribution both change over time (model with *predictable marks*, see, e.g., McNeil, Frey, and Embrechts, 2005) and the cases where the distribution of the exceedances is time constant (model with *unpredictable marks*, see, e.g., Chavez-Demoulin, 2005). The idea of both is a combination of the standard POT method and self-exciting (Hawkes) processes, see Hawkes (1971). To be specific, the rate of crossing the initial threshold u_i in (5) is modeled conditionally, i.e., $\tau_i(u_i) \to \tau_i(t, u_i)$, given information of the exceedances up to time t with the rate of a Hawkes process given by

$$\tau_i(t, u_i) = \tau_i + \psi_i \sum_{j:0<T_{i,j}<t} c_i\left(\tilde{X}_{i,j}\right) g_i\left(t - T_{i,j}\right), \tag{7}$$

where $\tau_i > 0$ and $\psi_i \geq 0$ are parameters. The variables $T_{i,1}, T_{i,2}, \ldots$ stand for the

times and $\tilde{X}_{i,1}, \tilde{X}_{i,2}, \ldots$ for the marks of the past exceedances. The path of $\tau_i(t, u_i)$ depends both on the timing of past events and on their marks. After an exceedance at time $T_{i,j}$ with mark $X_{i,j}$ the intensity jumps by $\psi_i c(\tilde{X}_{i,j})$ and, through function $g_i(\cdot)$, tends toward the constant level τ_i in the absence of an event. Later in the paper we use a normalization of the argument of the $c_i(\cdot)$ function by h_i, e.g., the 95th quantile of the marginal distribution to make the parameters of $c_i(\cdot)$ comparable over models and dimensions. Throughout the text we then use the notation

$$v_i^*(t) = \sum_{j:0<T_{i,j}<t} c_i\left(\frac{\tilde{X}_{i,j}}{h_i}\right) g_i\left(t - T_{i,j}\right). \tag{8}$$

Thus the decay function $g_i(\cdot)$ determines the rate how an influence of events decays in time, and the impact function $c_i(\cdot)$ determines the contribution of events to the conditional rate. The commonly used specification for $g_i(\cdot)$ is an exponential function

$$g_i(s) = e^{-\gamma_i s}, \qquad \gamma_i > 0. \tag{9}$$

For the impact function we suggest to use

$$c_i(x_i) = (x_i)^{\delta_i}, \qquad \delta_i \geq 0. \tag{10}$$

Combining (5) and (7) and adding the self-exciting component $v_i^*(t)$ also to the scale parameter β of the GPD, we obtain

$$\tau_i(t, x_i) = \tau_i(t, u_i)\left(1 + \xi_i \frac{x_i - u_i}{\beta_i + \alpha_i v_i^*(t)}\right)^{-1/\xi_i}, \qquad x_i \geq u_i, \tag{11}$$

the conditional rate of exceeding a high level x_i, given information of exceedances over u_i up to time t. Here the excesses over the threshold u_i are conditionally independent and distributed according to the GPD with shape parameter ξ_i and scale parameter $\beta_i + \alpha_i v^*(t)$. Since the history affects the distribution of the marks, this is the model with predictable marks. Setting $\alpha_i = 0$ we get the model with unpredictable marks. The exceedances then occur in time according to a non-homogeneous Poisson process

whereas the excesses over the threshold u_i are independent and identically distributed according to the GPD, see Chavez-Demoulin (2005).

Technically, $\tau_i(t, x_i)$ is the conditional rate of a point process $N_i(t, x_i)$ that counts number of exceedances of the level x_i by $X_{i,1}, X_{i,2}, \ldots, X_{i,t}$. Note that $N_i(t, x_i)$ is a continuous process, but we observe $X_{i,t}$ only at discrete times. Taking this discreteness into account, we can approximate the conditional probability that $X_{i,t+1}$ crosses the level $x_i \geq u_i$ given the history of past exceedances $H_{i,t}$ as the probability of at least one exceedance in period $(t, t+1]$. Formally,

$$P(X_{i,t+1} \geq x_i \mid H_{i,t}) \approx 1 - P\left(N_i(t+1, x_i) - N_i(t, x_i) = 0 \mid H_{i,t}\right) = 1 - \exp\left(-\int_t^{t+1} \tau_i(s, x_i) ds\right). \quad (12)$$

2.2. Bivariate Model

As a first generalization, we extend the univariate self exciting POT models to the bivariate case. We propose to model the arrival of joint exceedances with the conditional rate $\tau(t, x_1, x_2)$, which, similarly to (12), can be used for estimating the probability of joint exceedances as

$$P\left(X_{1,t+1} \geq x_1, X_{2,t+1} \geq x_2, \mid H_t\right) \approx 1 - \exp\left(-\int_t^{t+1} \tau(s, x_1, x_2) ds\right). \quad (13)$$

For the construction of $\tau(t, x_1, x_2)$ we consider the rate of joint exceedances $\tau(x_1, x_2)$ for independent realizations and then incorporate the self-exciting component.

For iid random vectors $\mathbf{X}_t = (X_{1,t}, X_{2,t})$, $t = 1, ..., n$ we assume that its distribution function F is in the domain of attraction of a multivariate extreme value distribution function G. Then there exist sequences $a_{i,n} > 0$ and $b_{i,n} \in \mathbb{R}$ such that for $x_{i,n} = a_{i,n} x_i + b_{i,n}$ it holds

$$\lim_{n \to \infty} nP(X_{1,t} > a_{1,n} x_1 + b_{1,n} \text{ or } X_{2,t} > a_{2,n} x_1 + b_{2,n}) = -\log G(x_1, x_2). \quad (14)$$

Since G is a multivariate extreme value distribution function, its marginal distributions G_i are univariate extreme value. We denote the shape parameter of G_i by ξ_i.

The point process

$$N_n(B) = \sum_{t=1}^{\infty} I_{\left\{\left(i/n,(1+\xi_1(X_{1,t}-b_{1,n})/a_{1,n})^{1/\xi_1},(1+\xi_2(X_{2,t}-b_{2,n})/a_{2,n})^{1/\xi_2}\right)\in B\right\}} \tag{15}$$

converges in distribution to a Poisson point process with rate $\lambda \times \nu$, where λ is the Lebesgue measure and ν the *exponent measure* [1]. For details see Theorem 6.1.11 in de Haan and Ferreira (2006). Considering the point process $N_n(B)$ it is convenient to transforms the margins of $\mathbf{X}_t$ to a unit Frechet distribution using $y = \exp(-1/x), x > 0$. $\mathbf{Y}_t = (Y_{1,t}, Y_{2,t})$ denotes a vector that has unit Frechet marginals and preserves the dependence structure of $\mathbf{X}_t = (X_{1,t}, X_{2,t})$, $t = 1, ..., n$. Due to the max-stability property of unit Frechet variables we can consider the process of points $(Y_{1,t}/n, Y_{2,t}/n)$ in representation (15).

For $\mathbf{Y}_t$ we have

$$\lim_{n\to\infty} nP(Y_{1,t} > ny_1, Y_{2,t} > ny_2) = 1/y_1 + 1/y_2 - V_\theta(y_1, y_2), \tag{16}$$

where $V_\theta(y_1, y_2)$ is the exponent measure of the set $\mathbb{R}^2_+ \setminus [(0, y_1) \times (0, y_2)]$ and it holds that $V_\theta(y_1, y_2) = -\log G(\frac{y_1^{\xi_1}-1}{\xi_1}, \frac{y_2^{\xi_2}-1}{\xi_2})$, while θ denotes dependence parameter of V. The right-hand side of (16) determines the average number of joint exceedances over (ny_1, ny_2) by n iid realizations of $\mathbf{Y}_t$ as $n \to \infty$. Assuming that this asymptotic argument holds for a finite sample size and exploiting the homogeneity property of the exponent measure, i.e., $V_\theta(ay_1, ay_2) = a^{-1}V_\theta(y_1, y_2)$ with $a > 0$, the rate of occurrence of joint exceedances over some high threshold (y_1, y_2) by a single realization of $\mathbf{Y}_t$ can be approximated as $1/y_1 + 1/y_2 - V_\theta(y_1, y_2)$. In terms of the initial vector $\mathbf{X}_t$

[1] There is an inherent relation between the exponent measure ν and an extreme value distribution G, namely, $G(x_1, x_2) = \exp(-\nu(\mathbb{R}^2_+ \setminus [(0, x_1) \times (0, x_2)]))$. The name of ν is due to this relation.

and the high threshold (x_1, x_2) this can be expressed as

$$\tau(x_1, x_2) = -\log(1 - \tau_1(x_1)) - \log(1 - \tau_2(x_2)) - V_\theta\left(\frac{1}{-\log(1 - \tau_1(x_1))}, \frac{1}{-\log(1 - \tau_2(x_2))}\right), \quad (17)$$

where we used the fact that over some high threshold u_i the tail of $X_{i,t}$ can approximated by the GPD, namely, $P(X_{i,t} > x_i) = \tau_i(x_i)$, see (6), and hence $Y_{i,t} \sim -1/\log(1 - \tau_i(X_{i,t}))$.

The rate (17) is derived for iid realizations of (X_1, X_2) and we add self-exciting components in order to capture possible temporal dependence between joint exceedances. We add it both to the processes of marginal exceedances and to the function V_θ obtaining the conditional rate of joint exceedances

$$\tau(t, x_1, x_2) = -\log(1 - \tau_1(t, x_1)) - \log(1 - \tau_2(t, x_2)) - V_{\theta(t)}\left(\frac{1}{-\log(1 - \tau_1(t, x_1))}, \frac{1}{-\log(1 - \tau_2(t, x_2))}\right), \quad (18)$$

where $V_{\theta(t)}$ denotes the exponent measure V_θ, but with a time-varying parameter. A time-varying structure of $V_{\theta(t)}$ accounts for temporal changes in the extremal dependence structure as, for example, during financial turmoil.

The processes of marginal exceedances are chosen as in the self-exciting POT model with predictable marks replacing $\tau_i(x_i)$ in (17) with $\tau_i(t, x_i)$ as in (11). This provides the mechanism how not only the joint exceedances, but also the marginal ones can contribute to occurrence of joint extremes. It accounts for the observation that shocks do not always occur simultaneously and it takes some time for the transmission to take place.

The parametric specification of $V_{\theta(t)}$ is still open and there are many parametric families of dependence structure in MEVT. With respect to applications, the dependence structure should both be as flexible as possible and be able to capture an asymmetric dependence structure, in the sense that $V_\theta(y_1, y_2) \neq V_\theta(y_2, y_1)$. This allows for asymmetric responses of the probability of joint exceedances to exceedances

of the individual variables implying interesting economic interpretations. For example, the stock market of a small country may react strongly to shocks to the US stock market, but not vice versa. We suggest to use the dependence function of the Gumbel copula[2]. It has a simple structure with only one parameter $\theta \geq 1$, which makes it easy to add the time-dependent part and to extend it to an asymmetric form. It also can be extended to dimensions beyond two, which is advantageous for the multivariate extension in Section 2.3. Furthermore, its dependence function in the tail is almost identical to the one of the t-copula for any choice of the parameters of the t-copula and is thus very flexible, see Demarta and McNeil (2005) for details.

With $V_\theta(y_1, y_2)$ of the asymmetric Gumbel copula

$$V_\theta(y_1, y_2) = \frac{(1-w_1)}{y_1} + \frac{(1-w_2)}{y_2} + \left(\left(\frac{w_1}{y_1} \right)^\theta + \left(\frac{w_2}{y_2} \right)^\theta \right)^{1/\theta},$$

the conditional rate of joint exceedance in (18) is given by

$$\begin{aligned} \tau(t, x_1, x_2) = w_1(-\log(1-\tau_1(t, x_1))) + w_2(-\log(1-\tau_2(t, x_2))) - \\ \left((-w_1 \log(1-\tau_1(t, x_1)))^{\theta(t)} + (-w_2 \log(1-\tau_2(t, x_2)))^{\theta(t)} \right)^{1/\theta(t)}. \end{aligned} \quad (19)$$

The parameters w_1 and w_2 account for asymmetry in the dependence structure. When $w_1 = w_2 = 1$ we obtain the symmetric Gumbel model. The time-varying dependence parameter $\theta(t)$ is parameterized as

$$\theta(t) = \theta_m + \psi_m v_m^*(t). \quad (20)$$

The self-exciting component $v_m^*(t)$ accounts for changes in the degree of the extreme dependence. Note that all components of the right-hand side of the above equation have a subscript m indicating that the parameters relate to the model of joint-exceedances, in contrast to the model of marginal exceedances. The structure of

[2]We initially also considered the Galambos copula, but its fit was inferior for all applications we considered.

$v_m^*(t)$ is the same as of $v_i^*(t)$ in the univariate self-exciting POT model in (8), that is

$$v_m^*(t) = \sum_{j:0<T_j<t} c_m \left(\frac{\tilde{X}_{1,j}}{h_1}, \frac{\tilde{X}_{2,j}}{h_2} \right) g_m(t - T_j), \tag{21}$$

where T_j stands for the time of joint exceedance. The function $c_m \left(\frac{\tilde{X}_{1,j}}{h_1}, \frac{\tilde{X}_{2,j}}{h_2} \right)$ is increasing in both arguments and determines the contribution of a joint exceedance with a mark $(\tilde{X}_{1,j}, \tilde{X}_{2,j})$ to the dependence parameter $\theta(t)$, whereas $g_m(\cdot)$ defines the rate how this contribution decays in time after the occurrence. As before, h_1 and h_2 are normalizing constants, for instance the 95th quantiles of the marginal distribution functions.

Again, $g_m(\cdot)$ and $c_m(\cdot)$ have to be specified. For the decay function $g_m(\cdot)$ we use the same form as for g_i in (9), namely

$$g_m(s) = e^{-\gamma_m s} \qquad \gamma_m > 0. \tag{22}$$

The choice of the impact function $c_m(\cdot)$ is ambiguous and may depend on the specific problem considered. We suggest the following specifications:

$$c_m(x_1, x_2) = \kappa_1 x_1 + \kappa_2 x_2, \tag{23}$$

$$c_m(x_1, x_2) = (\kappa_1 x_1 + \kappa_2 x_2)^{\delta_m}, \qquad \delta_m \geq 0, \tag{24}$$

$$c_m(x_1, x_2) = \max(x_1, x_2)^{\delta_m}, \qquad \delta_m \geq 0, \tag{25}$$

where $\kappa_1 \geq 0$ and $\kappa_2 \geq 0$ are weighting coefficients. The higher κ_i, the stronger the influence of a mark of the i-th margin on the dependence between the extreme events. This may lead to asymmetric responses to individual exceedances and the effect may possibly overlap with the effect captured by using an asymmetric copula. It is an empirical issue whether this full flexibility is necessary.

2.3. Multivariate Extension

In analogy with the bivariate case, we assume that the distribution function F of the vector $\mathbf{X} = (X_1, \ldots, X_d)$ is in the domain of attraction of the multivariate

extreme value distribution G. Hence, there exist $a_{i,n} > 0$ and $b_{i,n} \in \mathbb{R}$, for $i = 1, \ldots, d$, such that for $x_{i,n} = a_{i,n}x_i + b_{i,n}$ it holds that

$$\lim_{n\to\infty} n(1 - F(x_{1,n}, \ldots, x_{d,n})) = -\log G(x_1, x_2, \ldots, x_d), \tag{26}$$

see de Haan and Ferreira (2006). Transforming the margins of $\mathbf{X}$ to a unit Frechet distribution and denoting this vector by $\mathbf{Y} = (Y_1, \ldots, Y_d)$, the joint tail of $\mathbf{Y}$ can be expressed as

$$\lim_{n\to\infty} nP(\mathbf{Y} \geq \mathbf{y}n) = \sum_{\mathbf{i}\in\mathbf{D}^*} (-1)^{1+\sum_{k=1}^{d} i_k} V_\theta \left(y_1^{i_1}\infty^{1-i_1}, \ldots, y_d^{i_d}\infty^{1-i_d}\right), \tag{27}$$

where $\mathbf{i} = (i_1, \ldots, i_d)$, $\mathbf{D}^* = [0,1]^d \setminus \{0, \ldots, 0\}$, $\mathbf{y} = (y_1, \ldots, y_d)$, $V_\theta(y_1, \ldots, y_2)$ is the exponent measure of the set $\mathbb{R}_+^d \setminus [(0, y_1) \times \ldots \times (0, y_d)]$, and it holds that $V_\theta(y_1, \ldots, y_d) = -\log G(\frac{y_1^{\xi_1}-1}{\xi_1}, \ldots, \frac{y_d^{\xi_d}-1}{\xi_d})$. Note that ∞^{1-i_k} symbolically stands for 1, if $i_k = 1$, and for ∞, if $i_k = 0$. Defining the rate of joint exceedances in the same way as in Section 2.2 and adding the self-exciting component to the process of marginal exceedances and to the parameter of the exponent measure, the conditional rate of joint exceedances of the initial vector $\mathbf{X}_t$ over a high threshold $\mathbf{x} = (x_1, ..., x_2)$ can be expressed as

$$\tau(t, \mathbf{x}) = \sum_{\mathbf{i}\in\mathbf{D}^*} (-1)^{1+\sum_{k=1}^{d} i_k} V_{\theta(t)} \left(\frac{\infty^{1-i_1}}{(-\log(1 - \tau_1(t, x_1)))^{i_1}}, \ldots, \frac{\infty^{1-i_d}}{(-\log(1 - \tau_d(t, x_d)))^{i_1}} \right), \tag{28}$$

where $\tau_i(t, x_i)$ is as in (11) and $V_{\theta(t)}$ denotes the dependence function V_θ with a time-varying parameter. Joint exceedances of $\mathbf{X}_t$ over the level $\mathbf{x}$ by one realization of $\mathbf{X}$ can be modeled with a Poisson process with rate $\tau(t, \mathbf{x})$. In the same way as in the bivariate case, we interpret $\tau(t, \mathbf{x})$ as the conditional rate of joint exceedances by the vector $\mathbf{X}_t$ over the threshold $\mathbf{x}$, given the information of marginal and joint exceedances over the initial threshold $\mathbf{u} = (u_1, \ldots, u_d)$ up to time t.

For the non-exchangeable Gumbel copula, see Tawn (1990), V_θ is given by

$$V_\theta(y_1,\ldots,y_d) = \sum_{s\in S}\left\{\left(\sum_{i\in s} w_{i,s}/y_i\right)^{\theta_s}\right\}^{1/\theta_s}, \tag{29}$$

where S is the set of all non-empty subsets of $\{1,\ldots,d\}$ and the parameters are constrained by $\theta_s \geq 1$ for all $s \in S$, $w_{i,s} = 0$ if $i \notin s$, $w_{i,s} \geq 0$, $i = 1,\ldots,d$ and $\sum_{s\in S} w_{i,s} = 1$, see also Coles and Tawn (1991). V_θ in (29) is overparameterized for most applications, as it contains $2^{d-1}(d+2)-(2d+1)$ parameters. We suggest to use a shortened form of V_θ which is both concise and possesses an asymmetric structure:

$$V_\theta(y_1,\ldots,y_d) = \sum_{i=1}^{d}\left(\frac{1-w_i}{y_i}\right) + \left\{\sum_{i=1}^{d}\left(\frac{w_i}{y_i}\right)^{\theta}\right\}^{1/\theta}, \tag{30}$$

where $0 < w_k \leq 1$, for $k = 1,\ldots,d$, are the asymmetry parameters. If $w_k = 1$ for all $k = 1,\ldots,d$ we obtain the exponent measure of the exchangeable version of the Gumbel copula. With V_θ as above the conditional rate $\tau(t,\mathbf{x})$ takes the form

$$\tau(t,\mathbf{x}) = \sum_{\mathbf{i}\in\mathbf{D}^*}(-1)^{1+\sum_{k=1}^{d} i_k}\left[\left(\sum_{k=1}^{d}\left(-w_k i_k \log(1-\tau_k(t,x_k))\right)^{\theta(t)}\right)^{1/\theta(t)} + \sum_{k=1}^{d}(1-w_k)i_k(-\log(1-\tau_k(t,x_k)))\right]. \tag{31}$$

The time variation of the dependence parameter $\theta(t)$ is modeled the same way as for the bivariate model, see (20). The structure of $v_m^*(t)$ is given by

$$v_m^*(t) = \sum_{j:0<T_j<t} c_m\left(\frac{\tilde{X}_{1,j}}{h_1},\ldots,\frac{\tilde{X}_{d,j}}{h_d}\right) g_m(t-T_j), \tag{32}$$

where T_j stands for the time of joint exceedance, $\left(\tilde{X}_{1,j},\ldots,\tilde{X}_{d,j}\right)$ is the vector of marks of the joint extreme event, h_i a normalizing constant, $g_m(s) = \exp(-\gamma_m s)$,

$\gamma_m > 0$ the decay function, and $c_m(\cdot)$ the impact function. Analogously to the bivariate case we suggest the choices $c(x_1, \ldots, x_d) = \kappa_1 x_1 + \ldots + \kappa_d x_d$, $c_m(x_1, \ldots, x_d) = (\kappa_1 x_1 + \ldots + \kappa_d x_d)^{\delta_m}$, or $c_m(x_1, \ldots, x_d) = \max(x_1, \ldots, x_d)^{\delta_m}$, with $\delta_m \geq 0$, $\kappa_i \geq 0$, $i = 1, \ldots, d$.

3. Estimation, Goodness-of-Fit and Simulation

In this section derive the likelihood function for the estimation of our model (Section 3.1) and we discuss the goodness-of-fit (Section 3.2). Finally, in Section 3.3 we introduce a simulation algorithm.

3.1. Model Estimation

Statistical fitting of the model presented in Section 2 consists of estimating the marginal conditional rates $\tau_i(t, x_i)$, for $i = 1, \ldots, d$, and the conditional rate of joint exceedances $\tau(t, x_1, \ldots, x_d)$. In principle, one could estimate the model sequentially, but in order to gain statistical efficiency we maximize the log-likelihood functions of marginal and joint exceedances in one step, so we maximize the function

$$\log L = \log L_d^* + \sum_{k=1}^{d} \log L_i, \tag{33}$$

where L_i and L_d^* denote the likelihood functions of the marginal and joint rates of exceedances, respectively.

For convenience, we adopt the following notation: $\mathbf{X}_t = (X_{1,t}, \ldots, X_{d,t})$, $t = 1, 2, \ldots, n$ are the observations on which we estimate our model, $\mathbf{u} = (u_1, \ldots, u_d)$ is the initial threshold, i.e., u_i is a sufficiently high quantile of $X_{i,t}$; $T_{i,k}$ and $\tilde{X}_{i,k}$, with $k = 1, \ldots, N_{u_i}$, stand for the times and marks of the marginal exceedances over u_i and N_u is a set containing indexes of $\mathbf{X}_t$ such that if $t \in N_u$ then $\mathbf{X}_t$ exceeds $\mathbf{u}$ in at least one component.

The likelihood function for the marginal rate of exceedance, see McNeil, Frey,

and Embrechts (2005), is

$$L_i\left(\tau_i, \psi_i, \delta_i, \gamma_i, \xi_i, \beta_i, \alpha_i\right) = \exp\left(-n\tau_i - \psi_i \int_0^n v_i^*(s)ds\right) \prod_{j=1}^{N_{u_i}} \lambda_i\left(T_{i,j}, \tilde{X}_{i,j}\right), \quad (34)$$

where

$$\lambda_i(t, x_i) = \frac{\tau_i + \psi_i v_i^*(t)}{\beta_i + \alpha_i v_i^*(t)} \left(1 + \xi_i \frac{x_i - u_i}{\beta_i + \alpha_i v_i^*(t)}\right)^{-1/\xi_i - 1}$$

is the conditional intensity of the self-exciting POT model with predictable marks. The intensity $\lambda_i(t, x_i)$ is derived from

$$\int_{x_i}^{\infty} \lambda_i(t, s)ds = \tau_i(t, x_i).$$

With function v_i^* as in (8), the integral in (34) takes the form

$$\int_0^n v_i^*(s)ds = \int_0^n \sum_{j:0<T_{i,j}<s} e^{-\gamma_i(s-T_{i,j})} c_i\left(\frac{\tilde{X}_{i,j}}{h_i}\right) ds =$$
$$\frac{1}{\gamma_i} \sum_{k=1}^{N_{u_i}} \left[\left(e^{-\gamma_i T_{i,k}} - e^{-\gamma_i T_{i,k+1}}\right) \sum_{j=1}^{k} e^{\gamma_i T_{i,j}} c_i\left(\frac{\tilde{X}_{i,j}}{h_i}\right)\right],$$

with $T_{i,N_{u_i}+1} = n$.

Our approach to estimating the model of joint exceedances follows Coles and Tawn (1991). For iid random vectors $\mathbf{Y}_t = (Y_{1,t}, \ldots, Y_{d,t})$ that are in the domain of attraction of a multivariate extreme value distribution and have unit Frechet marginals, the likelihood for estimating the exponent measure V_θ of the set $A = \mathbf{R}_+^d \setminus \{(0, v_1) \times \ldots \times (0, v_d)\}$, where $v_1, \ldots, v_d$ are high thresholds of $\mathbf{Y}_t/n$, is as

$$L_A\left(\theta; \{\mathbf{Y}_t/n : t \in N_A\}\right) = \exp(-V_\theta(v_1, \ldots, v_d)) \prod_{t \in N_A} \mu(dr \times d\phi), \quad (35)$$

where θ denotes the parameter to be estimated, N_A contains indexes of the points $\mathbf{Y}_t/n$ that are in A, and $\mu(dr \times d\phi)$ denotes the intensity of occurrence of the points

$\mathbf{Y}_t/n$. The intensity is presented in pseudoradial, r_t, and angular, $\phi_t = (\phi_{1,t}, \ldots, \phi_{d,t})$, coordinates

$$r_t = \sum_{i=1}^{d} Y_{i,t}/n, \qquad \phi_{i,t} = Y_{i,t}/(nr_t), \qquad t = 1, \ldots, n, \qquad i = 1, \ldots, d, \tag{36}$$

for which it holds that $\mu(dr \times d\phi) = \frac{dr}{r^2} h(\phi) d\phi$, where $h(\phi)$ denotes a density of the spectral measure. Details about the spectral measure and its relationship to the exponent measure see, for example, Resnick (2007), de Haan and Ferreira (2006), or Coles and Tawn (1991).

To extend the above approach for estimation of our model we proceed by transforming the margins of $\mathbf{X}_t$ to a unit Frechet distribution (accounting for the time-varying features of exceedances of $\mathbf{X}_t$ over $\mathbf{u}$). The transformed vector is denoted by $\mathbf{Z}_t = (Z_{1,t}, \ldots, Z_{d,t})$, where

$$Z_{i,t} = \begin{cases} -1/\log\left(1 - \tau_i(t, X_{i,t})\right) & \text{if } X_{i,t} \geq u_i, \\ -1/\log(R(X_{i,t})/(n+1)) & \text{if } X_{i,t} < u_i. \end{cases} \tag{37}$$

In the above equation $\tau_i(t, x)$ denotes the distribution of the upper tail of the ith margin at time t, see (11), and $R(X_{i,t})$ denotes the rank of $X_{i,t}$. According to (37), the transformed threshold $\mathbf{u}$ is a time-varying vector $\mathbf{z}_t = (z_{1,t}, \ldots, z_{d,t})$, with $z_{i,t} = -1/\log\left(1 - \tau_i(t, u_i)\right)$. Incorporating (37) into (35) yields the likelihood function L_d^*, which can be used for parameter estimation of both the marginal models and the exponent measure. L_d^* is as follows

$$L_d^*(\theta_m, \phi_m, \gamma_m) = \exp\left(-\int_0^n V_{\theta(t)}(z_{1,t}, \ldots, z_{d,t})dt\right) \prod_{j \in N_u} \left(\frac{h(\phi_j)}{(nr_i)^{d+1}} \times \right.$$
$$\left. \prod_{\substack{i=1,\ldots,d: \\ Z_{i,j} > z_{i,j}}} \left[\frac{Z_{i,j}^2 \tau_i(j, u_i)^{-\xi_i}}{\beta_i + \alpha_i v_i^*(j)} \exp(1/Z_{i,j}) \left\{1 - \exp(-1/Z_{i,j})\right\}^{1+\xi_i}\right]\right), \tag{38}$$

where r_t and ϕ_t are pseudoradial and angular coordinates of $\mathbf{Z}_t$ given in (36). The

form of $V_{\theta(t)}$ is given in (29) and (30). The density of the spectral measure of the asymmetric Gumbel model suggested in (30), provided $\phi_t > 0$, is

$$h(\phi_t) = \left\{ \prod_{k=1}^{d-1} (k\theta(t) - 1) \right\} \left(\prod_{i=1}^{d} w_i \right)^{\theta(t)} \left(\prod_{i=1}^{d} \phi_{i,t} \right)^{-(1+\theta(t))} \left\{ \sum_{i=1}^{d} \left(\frac{w_i}{\phi_{i,t}} \right)^{\theta(t)} \right\}^{1/\theta(t)-d}. \tag{39}$$

Details for the likelihood function and the density of the spectral measure can be found in Coles and Tawn (1991). Note that the integral in (38) cannot be solved explicitly. In practice, this integral is approximated by a sum over all observations.

3.2. Goodness-of-fit

Applying the models of marginal and joint exceedances in practice, where true probabilities are unknown, it is vital to perform a goodness-of-fit procedure to check the performance of the model. One possible approach is to consider the *standardized inter-exceedances times*.

In the framework of Section 2.1, assume that a marginal exceedance over some level x_i occurs at time $T_{i,j}$. A time interval between this exceedance and the next one is random. Denote this random time interval by D. According to (12), conditioning on the information set $H_{i,T_{i,j}}$, the probability that in the next d periods at least one exceedance occurs is

$$P(D \leq d \mid H_{i,T_{i,j}}) = 1 - \exp\left(- \int_{T_{i,j}}^{T_{i,j}+d} \tau_i(s, x_i) ds \right).$$

Recalling that for a continuous random variable X with distribution function F, $F(X)$ is uniformly distributed on the unit interval, we obtain

$$\int_{T_{i,j}}^{T_{i,j}+L} \tau_i(s, x_i) ds = -\log U, \tag{40}$$

where $U \sim U[0,1]$. It follows from (40), that for times $T_{i,j}$, $T_{i,j+1}$ of consecutive marginal exceedances the value $\int_{T_{i,j}}^{T_{i,j+1}} \tau_i(s, x_i) ds$ is a realization of an exponential random variable with mean 1.

With the analogous consideration as for the marginal exceedances we obtain a similar result for the joint exceedances. For times $T_{i,j}$, $T_{i,j+1}$ of consecutive joint exceedances the value $\int_{T_{i,j}}^{T_{i,j+1}} \tau(s, x_1, ..., x_d)ds$ is a realization from a standard exponential distribution.

Assuming that we have observed n points of marginal or joint exceedances at times T_j, $j = 1, 2, \ldots, N$, we construct the goodness-of-fit procedure as follows. Define the sequence of the *standardized inter-exceedance times* $(\chi_1, \ldots, \chi_{N-1})$ by:

$$\chi_j = \int_{T_j}^{T_{j+1}} \tau(s)ds, \qquad j = 1, 2, \ldots, N-1$$

where $\tau(s)$ denotes the conditional rate of marginal or joint exceedances. It is apparent that the χ_j theoretically follow a standard exponential distribution. In practice, this will be only an approximate result, because the true values of the parameters are unknown. The goodness-of-fit can be checked either graphically using QQplots, or using a formal goodness-of-fit test such as the Kolmogorov-Smirnov or Anderson-Darling tests to test whether the estimates of χ_j follow a standard exponential distribution. As part of a goodness-of-fit procedure one can also analyze the ability of random data simulated from the model to reproduce certain characteristics of the data. The simulation procedure is described in the next section.

3.3. Simulation

Simulating a sample with n observations from the bivariate model of joint exceedances, for $t \in \{1, 2, \ldots, n\}$ we model the occurrence of exceedances over the initial threshold as a discrete random variable with the following possible outcomes:

i) exceedance by only the 1st margin, with probability

$$P(X_{1,t} \geq u_1, X_{2,t} < u_2 \mid H_{t-1}) = \exp\left(-\int_{t-1}^{t} \tau(s, u_1, u_2)ds\right) - \exp\left(-\int_{t-1}^{t} \tau_1(s, u_1)ds\right),$$

ii) exceedance by only the 2nd margin, with probability

$$P(X_{1,t} < u_1, X_{2,t} \geq u_2 \mid H_{t-1}) = \exp\left(-\int_{t-1}^{t} \tau(s, u_1, u_2)ds\right) - \exp\left(-\int_{t-1}^{t} \tau_2(s, u_2)ds\right),$$

iii) exceedance by both the 1st and 2nd margins, with probability

$$P(X_{1,t} \geq u_1, X_{2,t} \geq u_2 \mid H_{t-1}) = 1 - \exp\left(-\int_{t-1}^{t} \tau(s, u_1, u_2)ds\right),$$

iv) no exceedance at *all*, with probability

$$P(X_{1,t} < u_1, X_{2,t} < u_2 \mid H_{t-1}) = \exp\left(-\int_{t-1}^{t} \tau_1(s, u_1)ds\right) + \exp\left(-\int_{t-1}^{t} \tau_2(s, u_2)ds\right) - \exp\left(-\int_{t-1}^{t} \tau(s, u_1, u_2)ds\right).$$

The marks of exceedances follow a GPD with shape parameter ξ_i and a time varying scale parameter $\beta_i + \alpha_i v_i(t)$, i.e., the marks $\tilde{x}_{i,t}$ can be generated from $\tilde{X}_{i,t}$, where

$$\tilde{X}_{i,t} = u_i + \frac{\beta_i + \alpha_i v_i^*(t)}{\xi_i}(U^{-\xi_i} - 1), \qquad U \sim U[0, 1]. \tag{41}$$

By updating the rates of marginal and joint exceedances, as described in Section 2.2, and repeating the above simulation procedure for each $t \in \{1, 2, \ldots, n\}$ we obtain a path of simulated exceedances. Simulating the data it may happen that for some t, $\tau_i(t, u_i) \geq 1$, which is in contradiction with (6) and may lead to explosive behavior of the simulated path. To avoid this, $\tau_i(t, u_i)$ is to be bounded from above by 1, i.e. $\tau_i(t, u_i) = \min(\tau_i + \psi_i v_i^*(t), 1)$. In Section 4.3.1 we illustrate the use of the simulation algorithm.

4. Application to Financial Data

In this section we apply the model of joint exceedances to financial data. We consider a two- and a four-dimensional application of the model, and focus on describing the behavior of extreme (negative) returns in financial markets worldwide and in the European banking sector.

4.1. Data and Preliminary Analysis

In order to illustrate the two dimensional model we consider extreme negative returns in European and the US financial markets. The data consists of daily log-returns of the two Morgan Stanley Capital International (MSCI) indexes for US (MSCI-USA) and Europe (MSCI-EU) for the period January 1, 1990 to January 13, 2012, resulting in 5749 observations. The MSCI-USA index is designed to measure large and mid cap equity performance of the US equity market, whereas the MSCI Europe Index measures the equity performance of the developed markets in Europe and consists of the following country indices: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom[3].

The four-dimensional model is applied to joint extreme negative returns on equity of four major European banks, namely Royal Bank of Scotland (RBS), United Bank of Switzerland (UBS), Deutsche Bank (DB), and HSBC Holdings (HSBC), embracing the period October 20, 1993 to January 13, 2012. The sample consists of 4768 observations. The summary statistics for all series can be found in Table 1. All six returns series show high excess kurtosis suggesting Frechet type tails for the returns distribution. Note that the extremely high skewness and kurtosis for RBS is the effect of two extremely large negative returns. We decided not to delete these observations since the methods we apply have a certain robustness to outliers and we are in fact interested in very extreme events. Furthermore, since the estimate of the tail index of left tail of RBS return time series (see Table 5) is larger than 0.25 suggesting that kurtosis does not exist for this time series. Note that for estimation of our model we use negated daily log-returns on the equity, allowing us to look at the upper rather than the lower tail.

4.2. Extreme Value Condition

Recall that the multivariate extreme value condition is an underlying assumption of the model of joint exceedances presented in this paper. This assumption includes

[3]see www.msci.com for details

Table 1: Summary statistics

	MSCI-USA	MSCI-EU	DB	HSBC	RBS	UBS
Mean(%)	0.0232	0.0155	-0.0025	0.0192	-0.0246	-0.0100
St.Deviation	0.0116	0.0126	0.0234	0.0190	0.0330	0.0231
Skewness	-0.2468	-0.1844	0.1968	-0.2152	-8.2565	0.1289
Excess Kurtosis	8.7890	8.0630	8.9337	8.6815	272.8115	12.5427

both the marginal extreme value condition and the existence of extreme dependence. Thus, to be able to apply our model to real data it is crucial to test this assumption first.

We rely on the mean-excess function to verify if a GPD is an appropriate model for the excesses and to choose the initial threshold. Details on this and other methods may be found, e.g., in McNeil, Frey, and Embrechts (2005), Embrechts, Klueppelberg, and Mikosch (1997) or Resnick and Starica (1995), discussions about choosing a threshold in Chavez-Demoulin and Embrechts (2011). For positive-valued data $X_1, X_2, \ldots, X_n$ we estimate the mean-excess function as

$$e_n(v) = \frac{\sum_{i=1}^{n}(X_i - v)I_{\{X_i>v\}}}{\sum_{i=1}^{n} I_{\{X_i>v\}}}. \tag{42}$$

Plotting $\{X_{i,n}, e_n(X_{i,n})\}$, where $X_{i,n}$ denotes the ith order statistic, we consider a shape of the mean-excess function. If the shape looks approximately linear then this suggests that a GPD is an appropriate model for the excesses. The point where the mean-excess function visually becomes close to linear can be set as a threshold for GPD estimation. Figures 2 plots the estimates of mean-excess function for the first 6% of the sample upper order statistics. For all series, the GPD seems appropriate. We set the marginal initial threshold at points where the mean-excess function visually becomes (close to) linear (denoted by a solid vertical line in Figure 2). For MSCI-USA series the initial threshold is 2.4922% (97.7% quantile), for MSCI-EU is 2.86% (97.7% quantile), which results in 132 marginal exceedances for the two indexes, and in 53 joint exceedances. For DB the series the initial threshold is 4.2462% (96.22% quantile), for HSBC is 3.281% (96.22% quantile), for RBS is

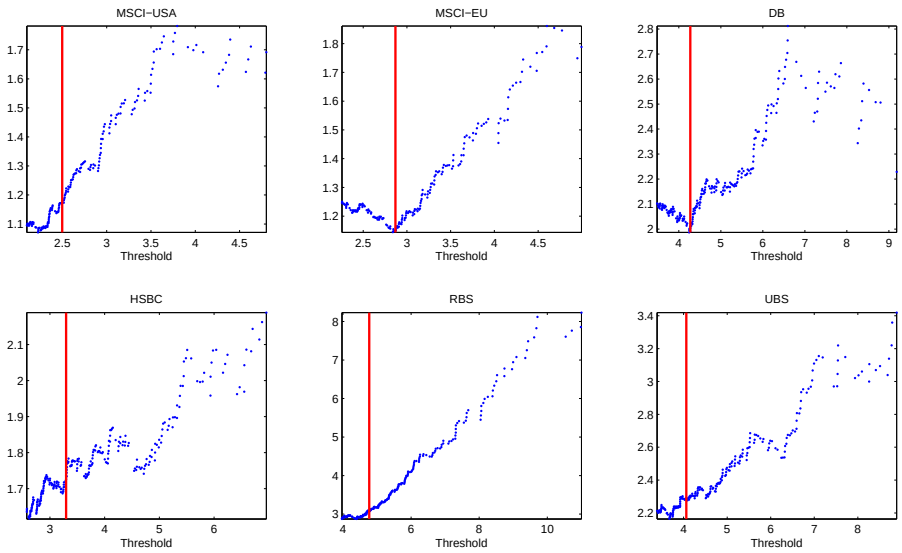

Figure 2: Sample mean excess plots of negated daily log-returns of the MSCI-USA, MSCI-EU, DB, HSBC, RBS, and UBS.

4.7179% (96.22% quantile), and for UBS is 4.0338% (96.22% quantile), with such thresholds there are 180 marginal (for all four indexes) and 32 joint exceedances.

Furthermore, since the multivariate extreme value condition is an underlying assumption of the model of joint exceedances, it is crucial to know whether there is asymptotic dependence in the data used for the model estimation. For this purpose, we employ a graphical test called a Q-curve, see de Haan and Ferreira (2006). Flat Q-curves indicate asymptotic independence. Figure 3 illustrates the Q-curve of the negative log returns of MSCI-USA and MSCI-EU, estimated on different number (k) of upper order statistics of the return series. The line labeled "*ind*" indicates the Q-curve in the independence case. The curves on Figure 3 differ significantly from a straight line indicating that there is no asymptotic independence between negative log returns of MSCI-USA and MSCI-EU indexes.

To visualize the Q-curve in the four-dimensional case, we report its three-dimensional projections. Analogously to the bivariate case, a flat convex shape of the Q-curve in the three dimensional case indicates the presence of asymptotic independence. If

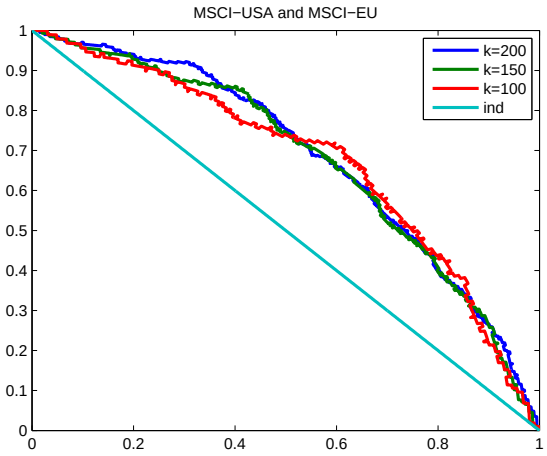

Figure 3: Estimated Q-curves on negated log-returns of MSCI-USA and MSCI-EU: $k = 100,\ 150,\ 200$.

the shape is concave, one expects no asymptotic independence. Figure 4 illustrates the three dimensional Q-curves of DB, HSBC, RBS, UBS return series, estimated on 200 upper order statistics of the return series. The curves on Figure 4 seem to differ significantly from a flat curve indicating that there is no asymptotic independence between negative log returns of DB, HSBC, RBS, UBS equity prices.

4.3. Applying the Model

We now present the estimation results for the two- and four-dimensional models of joint exceedances. In preliminary estimations we considered the three different specifications for the impact function $c(\cdot)$ we proposed Section 2.3 and found that none of the impact functions is distinguishable from a constant one, i.e., setting $c \equiv 1$. This implies that magnitudes of exceedances are irrelevant for modeling of the dependence parameter in our data sets. Hence, in the following we consider the models of joint exceedances with a constant impact function. The normalizing constants h_i introduced in equations (7), (21) and (32) are set equal to the 95% quantiles of the marginal distribution functions.

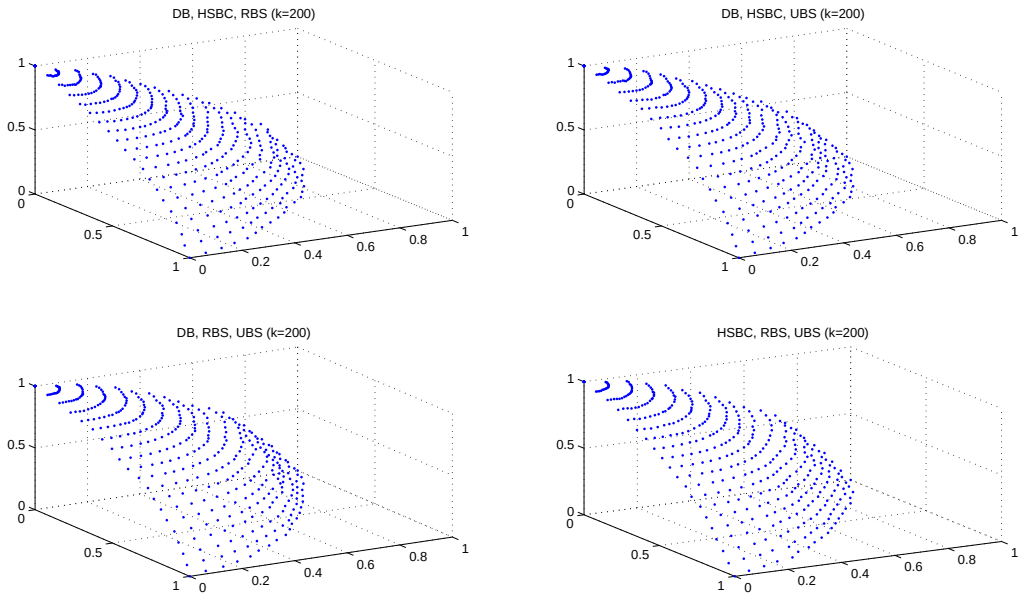

Figure 4: Estimated Q-curves on negated log-returns of DB, HSBC, RBS, and UBS.

4.3.1. Two-dimensional Model

In this section, we focus on joint extreme negative returns in Europe and USA. We use negated daily log-returns of MSCI-USA and MSCI-Europe indexes for estimation. Estimating the two-dimensional model of joint exceedances with the *asymmetric* Gumbel copula we conduct two likelihood tests on the following hypothesis: H_0 : $\psi_m = 0$, i.e., the dependence parameter $\theta(t)$ is constant, and $H_0 : w_1 = w_2 = 1$ that the dependence structure is symmetric. It turns out that there is significant evidence against the null hypotheses $H_0 : \psi_m = 0$ with a p-value of 0.00001, and that there is no significant evidence against the null hypothesis $H_0 : w_1 = w_2 = 1$ (p-values 0.9999). The parameter estimates along with 95% confidence intervals are reported in Tables 2-3. The confidence intervals are computed by using the profile log-likelihood

Table 2: Parameter estimates of the self-exciting POT model with predictable marks: two-dimensional model.

	MSCI-USA			MSCI-EU		
Parameter	Estimate	Conf. interval		Estimate	Conf. interval	
τ_i	0.0089	[0.0066	0.0115]	0.0079	[0.0059	0.0101]
ψ_i	0.0181	[0.0112	0.0283]	0.0215	[0.0136	0.0332]
γ_i	0.0434	[0.0327	0.0577]	0.0587	[0.0459	0.0746]
δ_i	0.7979	[0.3329	1.2420]	0.9351	[0.4929	1.3515]
ξ_i	0.1988	[0.0595	0.3692]	0.2016	[0.0770	0.3586]
β_i	0.5360	[0.4027	0.6972]	0.5272	[0.3933	0.6935]
α_i	0.1120	[0.0597	0.2011]	0.1214	[0.0654	0.2125]

Table 3: Parameter estimates of the model of joint exceedances: two-dimensional model.

Parameter	Estimate	Conf. interval	
θ_m	1.3793	[1.3029	1.4655]
ψ_m	0.1094	[0.0552	0.2080]
γ_m	0.0087	[0.0032	0.0175]

function, see, e.g., Coles (2001). Simulations and practical experience suggest these intervals provide better results than those derived by using the numerical Hessian matrix.

The estimates of the tail index suggest that the left part of the distribution of the MSCI-USA and MSCI-EU returns is quite heavy tailed, indicating substantial downward potential of the financial markets. Furthermore, comparing estimates of δ_i indicates that the marks of MSCI-EU exceedances over the initial threshold have a larger influence on the conditional rate of marginal exceedances than MSCI-USA exceedances have. The conditional rates of exceeding the initial threshold are reported in Figure 5. Figures 6 plots the time varying dependence parameter $\theta(t)$

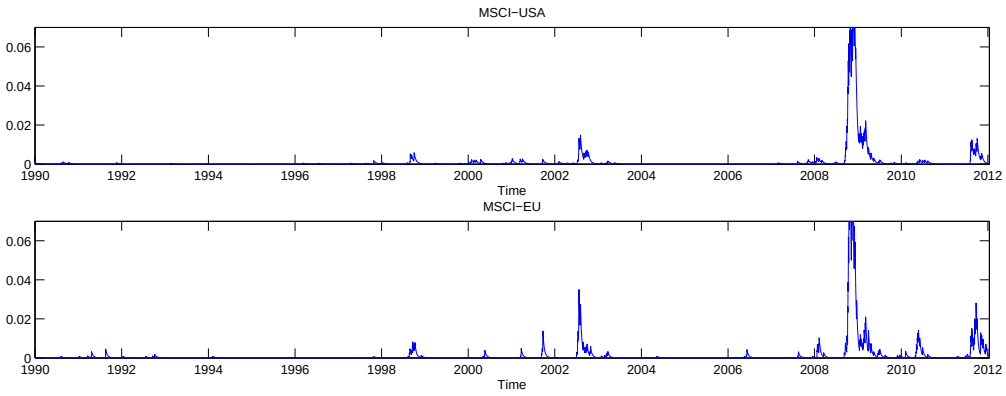

Figure 5: The estimated conditional rates of the marginal exceedances over the initial thresholds in the self-exciting POT model with predictable marks for negated log-returns of MSCI-USA and MSCI-EU indices.

(left panel) and the estimated conditional probabilities of joint exceedances over the initial threshold compared with the (constant) empirical probability of joint exceedances (right panel).

The model of joint exceedances provides a specific feedback mechanism describing how the mark of an exceedance affects the arrival of future events. Figure 7 illustrates the effect of different values of MSCI-USA negated returns at time $n+1$, on the components of the model at time $n+2$, where $n = 5749$ – number of observations used for estimation of the bivariate model. Note that we denote MSCI-USA's (MSCI-Europe's) negated return at time t by $X_{1,t}$ ($X_{2,t}$). Realizations of $X_{1,n+1}$ that are smaller than the threshold $u_1 = 2.49\%$ do not influence any component of the model. For $X_{1,n+1} = u_1$, however, the rate of marginal exceedance $\tau_1(n+2, u_1)$

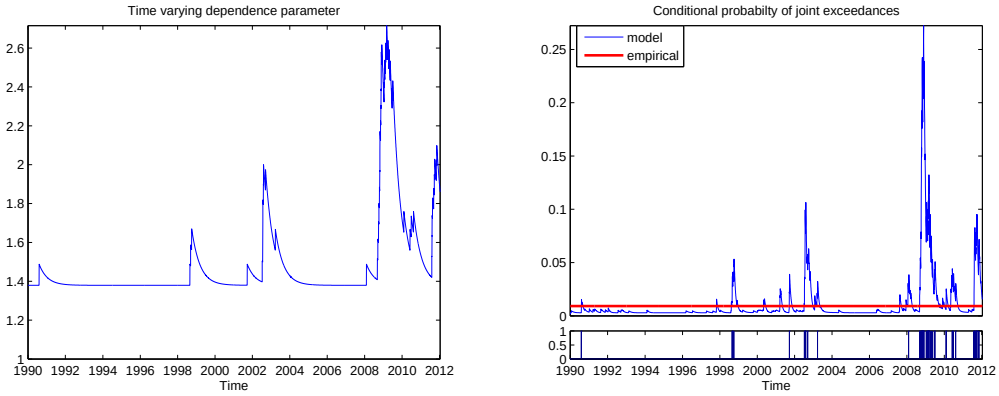

Figure 6: The estimated time-varying dependence parameter (left-hand panel) and the conditional probability of joint exceedances over the initial threshold (right-hand panel) in the two dimensional model. An indicator of times of joint exceedances (right-bottom panel).

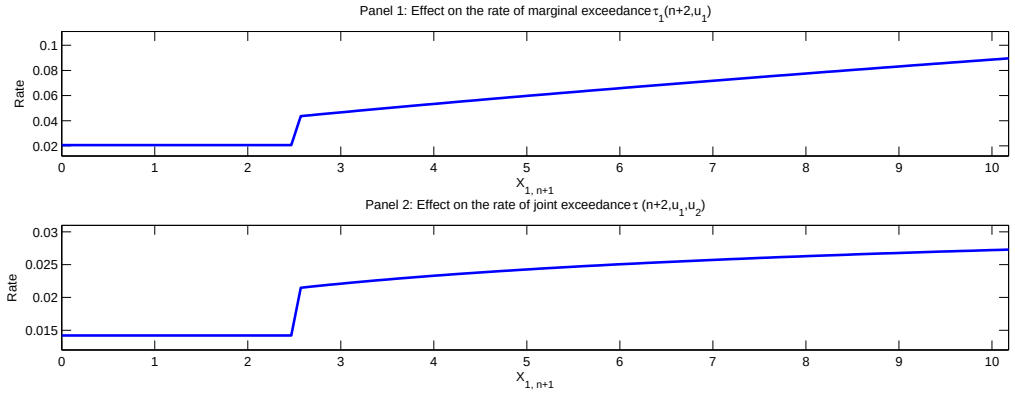

Figure 7: Effects of different values of $X_{1,n+1}$ on the rates of marginal and joint exceedances over the initial threshold at time $(n+2)$.

jumps (see Panel 1 of Figure 7) and increases smoothly for $X_{1,n+1} > u_1$, leading to a corresponding increase, through the relation (19), in the rate of joint exceedance $\tau(n+2, u_1, u_2)$ (Panel 2 of Figure 7). Assuming that there is no exceedance of $X_{2,n+1}$ over $u_2 = 2.86\%$, the dependence parameter $\theta(n+2)$ and the rate $\tau_2(n+2, u_2)$ are unaffected by the marginal exceedances of $X_{1,n+1}$. It is, of course, not desirable that the effect increases so strongly at a arbitrary threshold. However, this is the nature of the POT model and one is usually interested in more sizable exceedances.

In case of a joint exceedance at time $n+1$, we observe an increase in all components of the model, that is in $\tau_1(n+2, u_1)$, $\tau_2(n+2, u_2)$, and $\theta(n+2)$, that results in an increase of $\tau(n+2, u_1, u_2)$. Figure 8 illustrates the effect of different values of $X_{1,n+1}$ and $X_{2,n+1}$ at time $n+1$, on the rate of joint exceedances at time $n+2$. It is

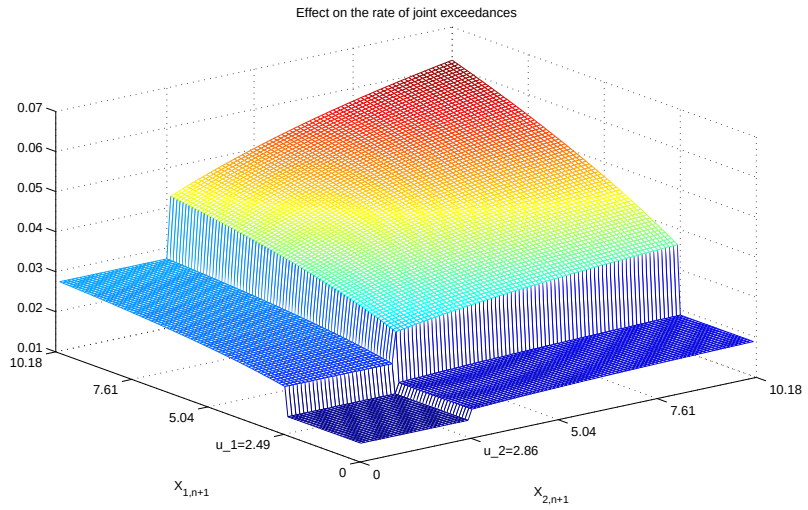

Figure 8: Effects of different values of $X_{1,n+1}$ and $X_{2,n+1}$ on the rate of joint exceedances over the initial threshold at time $(n+2)$.

important to be aware of the fact that the symmetrical dependence implied by our model does not mean that the marks of exceedances, standardized by h_i, have the same effect on the conditional rate of joint exceedances. This effect depends on the history of processes and, hence, it is time-varying. Figure 8 illustrates the asymmetry of influence of marks of exceedances on the conditional rate of joint exceedances over

the initial threshold.

Figure 9 shows the QQplot of the standardized inter-exceedances times against the standard exponential distribution and the bar plot of real and standardized times of joint exceedances. From the bar plots in Figures 9 (right panel), it is visible that the clusters in the original series of joint exceedances have been captured reasonably well indicating a good model fit. We have also added to the QQplot on the left-

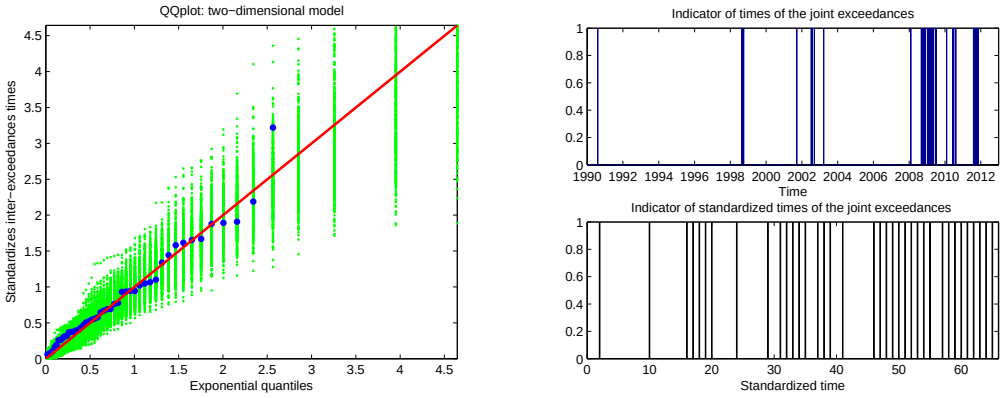

Figure 9: QQ-plot of the standardized inter-exceedances times vs standard exponential quantiles (left-hand panel). Bar plot indicating the real and standardized times of exceedances (right-hand panel). Two-dimensional model.

hand panel the analysis of 500 realizations of the exponential random variable (in green color), in order to illustrate what type of deviations one can expect. The standardized inter-exceedances times do not deviate very much from the standard exponential distribution suggesting the theoretical consistency of the model.

We now use the estimated parameters of Tables 2-3 to simulate sample paths as described in Section 3.3. Figure 10 plots simulated exceedances. The real exceedances of the data used for parameter estimation is shown in Figure 1.

Table 4 provides a comparison of the characteristics of simulated data, averaged of 1000 random samples of the same length as the data from the empirical application, with the characteristics of this data. Note that margin 1 corresponds to the MSCI-USA index and margin 2 to the MSCI-EU index. The results suggest that the simulated data captures the characteristics of the real data quite well, indicating

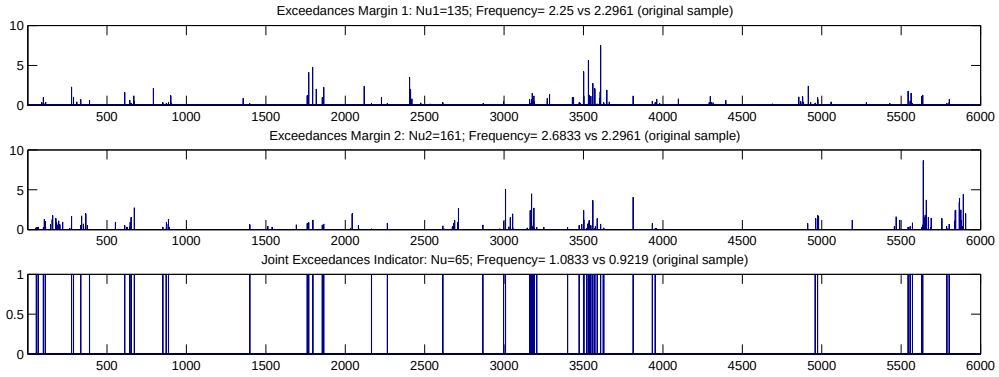

Figure 10: Simulated exceedances for the 1st margin (top panel) and the 2nd margin (middle panel). An indicator of times of joint exceedances (bottom panel).

that both the model specification is appropriate and simulation algorithm works.

Table 4: Comparison of the simulated data to the data used for the estimation

	Mean	std	Original sample
frequency of exceedances of margin 1	0.0265	0.0081	0.0230
frequency of exceedances of margin 2	0.0242	0.0093	0.0230
frequency of joint exceedances	0.0087	0.0031	0.0092
average mark of exceedances of margin 1	3.5516	0.3523	3.6622
average mark of exceedances of margin 2	3.9925	0.5648	4.0084

4.3.2. Four-dimensional Model

In this Section, we focus on joint extreme negative returns of equity of four major European banks: Deutsche Bank (DB), HSBC Holdings (HSBC), Royal Bank of Scotland (RBS), and UBS. As before, after estimating the four-dimensional model of joint exceedances with the *asymmetric* Gumbel copula we conduct two likelihood tests on the null hypotheses that the dependence parameter is constant and that the dependence structure is symmetric. We reject the first hypotheses, but not the second one with p-values of 0.0001 and 0.99, respectively. Parameter estimates along with 95% confidence intervals can be found in Tables 5-6. Comparing the estimates of δ_i and ψ_i parameters, we note that although the initial threshold was set on the same

Table 5: Parameter estimates of the self-exciting POT model with predictable marks: four-dimensional model.

	DB			HSBC		
Parameter	Estimate	Conf. Interval		Estimate	Conf. Interval	
τ_i	0.0143	[0.0086	0.0186]	0.0107	[0.0036	0.0215]
ψ_i	0.0190	[0.0115	0.0323]	0.0199	[0.0066	0.0398]
γ_i	0.0528	[0.0316	0.0676]	0.0286	[0.0137	0.0571]
δ_i	1.2259	[0.2268	1.5578]	0.5173	[0.1832	1.0185]
ξ_i	0.1811	[0.0605	0.3621]	0.1354	[0.0479	0.2665]
β_i	1.0710	[0.6996	1.4853]	1.0483	[0.4805	1.9001]
α_i	0.1149	[0.0534	0.2088]	0.1398	[0.0495	0.2751]
	RBS			UBS		
Parameter	Estimate	Conf. Interval		Estimate	Conf. Interval	
τ_i	0.0088	[0.0031	0.0173]	0.0106	[0.0037	0.0195]
ψ_i	0.0194	[0.0081	0.0365]	0.0223	[0.0093	0.0438]
γ_i	0.0251	[0.0104	0.0447]	0.0371	[0.0186	0.0615]
δ_i	0.3002	[0.1063	0.5911]	0.4981	[0.1764	0.9807]
ξ_i	0.3277	[0.1161	0.6452]	0.1017	[0.0360	0.2003]
β_i	0.7171	[0.2540	1.4118]	1.0675	[0.5337	1.9014]
α_i	0.2693	[0.0954	0.5302]	0.1980	[0.0701	0.3898]

Table 6: Parameter estimates of the model of joint exceedances: four-dimensional model.

Parameter	Estimate	Conf. Interval	
θ_m	1.4364	[1.2726	1.5164]
ψ_m	0.0254	[0.0126	0.0508]
γ_m	0.0119	[0.0065	0.0237]

level (96.22% quantile) for all indexes, the marks of exceedances have sufficiently different influence on the conditional rates of joint exceedances, with DB stocks being the most influencing factor. On the other hand, the estimates of γ_i show that the influence of DB exceedances on the marginal (joint) rates decay faster in time compared to other banks. This interplay between the impact and decay functions determines another mechanism of asymmetric responses of marginal events on the rate of joint exceedances. The conditional rates of exceeding the initial threshold are reported in Figure 11.

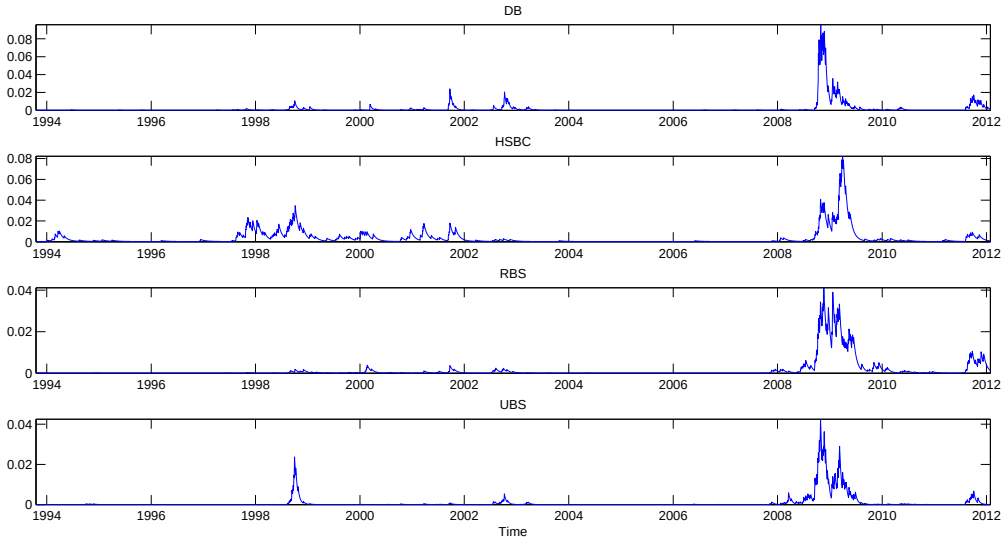

Figure 11: The estimated conditional rates of the marginal exceedances over the initial threshold in the self-exciting POT model with predictable marks for negated log-returns of DB, HSBC, RBS, and UBS stocks.

The time varying dependence parameter $\theta(t)$ and the estimated conditional probabilities of joint exceedances are depicted in Figure 12. Note that the dependence parameter in the two-dimensional model is more volatile than the one in the four-dimensional case, which can be explained by the estimates of ψ_m and γ_m. In the two-dimensional model the contribution of a joint exceedance to the dependence parameter is larger in magnitude (due to larger ψ_m) and decays slower in time (due

to the smaller γ_m) than in the four-dimensional case suggesting that shocks in the banking sector are less lasting and less severe than their overall influence on the European and US financial markets.

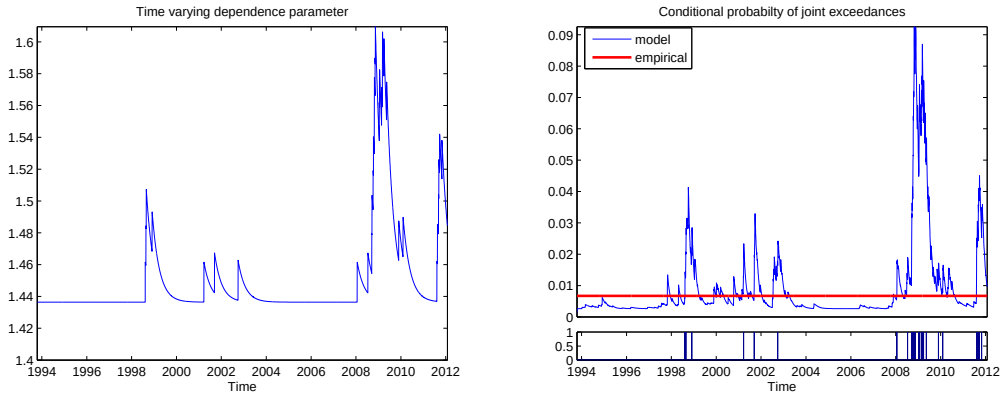

Figure 12: The estimated time-varying dependence parameter (left-hand panel) and the conditional probability of joint exceedances (right-hand panel) in the four-dimensional model. An indicator of times of joint exceedances (right-bottom panel).

The results for the goodness-of-fit can be found in Figure 13. Again, the fit looks satisfactory and the clusters of joint exceedances have been captured reasonably well. As before, in order to illustrate reasonable deviations from a theoretical case, we have added the analysis of 500 realizations of the exponential random variable (in green color) to the QQplot on the left-hand panel of Figure 13. Although the standardized inter-exceedances times do not deviate too much from a standard exponential distribution, the fit is not perfect, which may partly be caused the bias due to the estimation of 31 parameters. Another factor influencing the fit of the multivariate model of joint exceedances is the fit of the marginal models, namely, the self-exciting POT with predictable marks. Figure 14 illustrates the QQplot of the marginal standardized inter-exceedances times against the standard exponential distribution. Figure 15 provides the QQplot of marks of the marginal exceedances transformed, by the GPD with the appropriate shape ξ_i and scale parameter $\beta_i + \alpha_i v_i^*(t)$, to a standard exponential distribution. The fairly good fit of the marginal models illustrates a *structural assumption* of POT-models, i.e., conditioning over the threshold

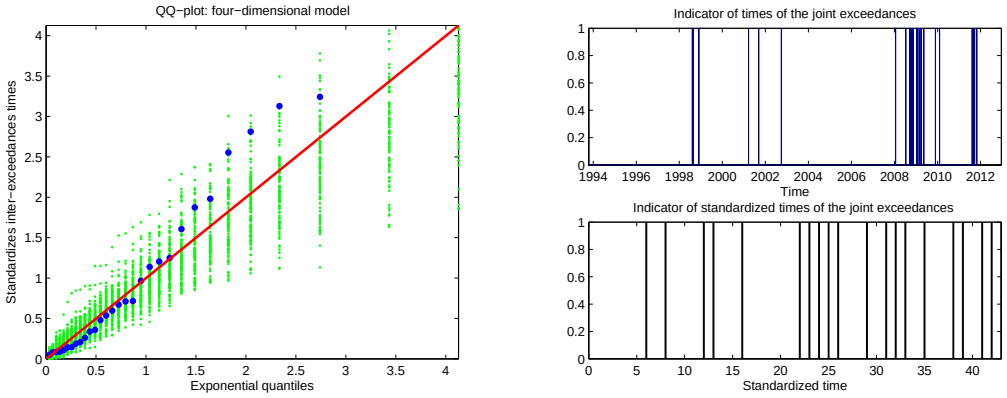

Figure 13: QQ-plot of the standardized inter-exceedances times vs standard exponential quantiles (left-hand panel). Bar plot indicating the real and standardized times of exceedances (right-hand panel). Four-dimensional model.

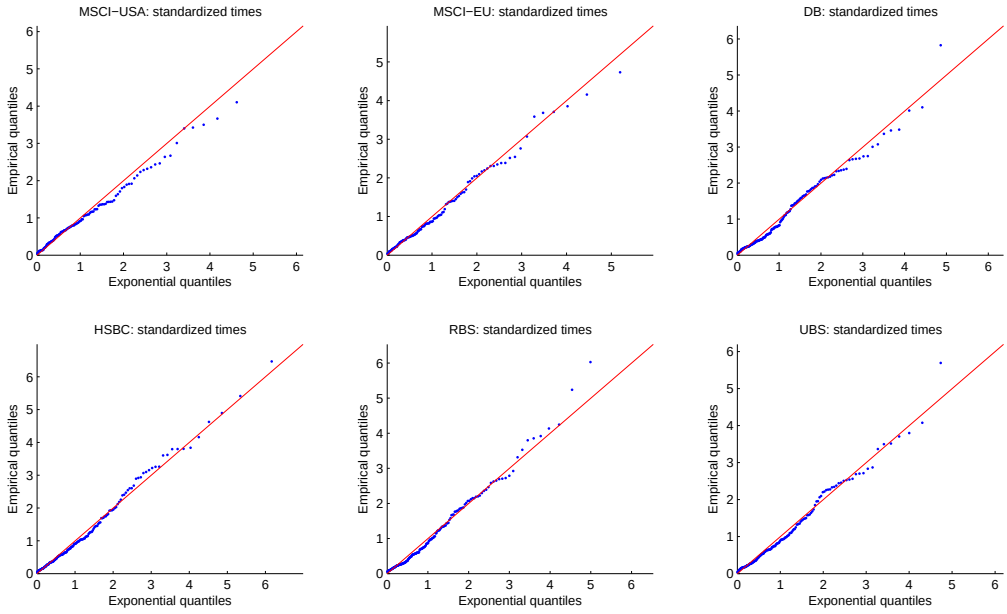

Figure 14: QQ-plot of the marginal standardized inter-exceedances times vs standard exponential quantiles.

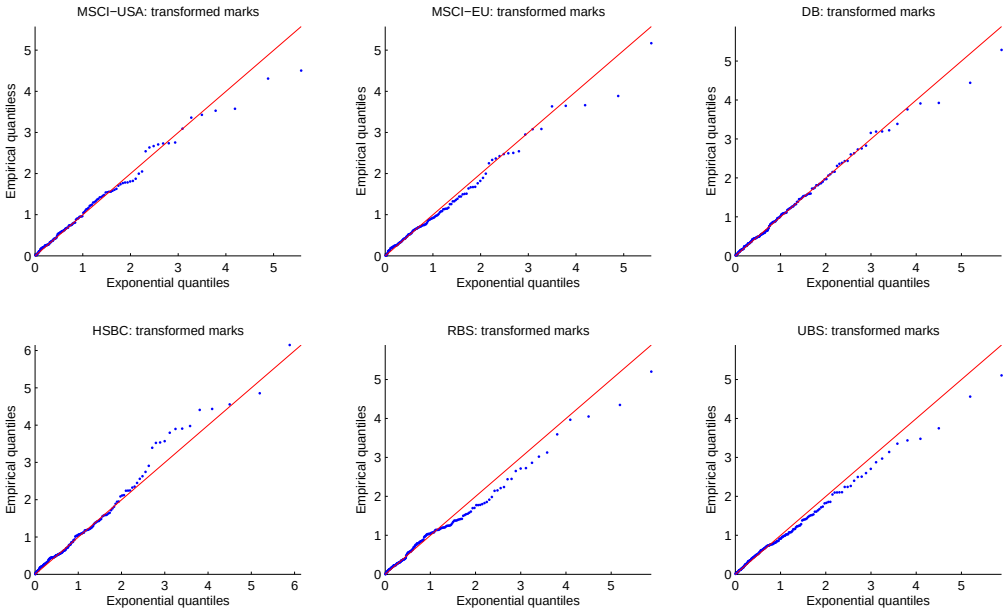

Figure 15: QQ-plot of the transformed marks vs standard exponential quantiles.

that is higher than the initial one, the GPD with the same shape parameter is still appropriate for the tail modeling.

The fit of the model of joint exceedances in the two and four-dimensional cases suggest that the model provides an efficient way to quantify the effects that cause the clustering of extreme financial returns. Among others, these effects are the reaction of markets to common economic factors and interplay between markets through time-varying linkages. Note that being able to quantify the effects our model cannot explain the source of clustering and contagion, because our model is decidedly reduced-form. To show when and where exactly the shock occurs, a pure qualitative analysis is required.

5. Conclusion

In the present paper, we develop a multivariate approach to model joint exceedances over high thresholds by using an extreme value distribution and a self-exciting processes as a way to capture the temporal dependence in extreme events. In our model both marginal exceedances of certain dimensions, as well as joint exceedances may affect the probability of joint extremes to a different extent. The model captures the feedback from marginal extreme events as well as changes in the dependence intensity between extreme events.

We apply the model to equity data, studying the patterns of self-exciting feature in the financial markets (USA, Europe) and in the European banking sector (Deutsche Bank, RBS, HSBC, and UBS). Goodness-of-fit results demonstrate a reasonable fit of the model and suggest the empirical importance of the self-exciting feature for modeling marginal extreme events and for capturing changes in the interdependencies between extreme events. We find that the full flexibility of our model is not required to describe that data, as the responses to shocks and the dependence between the exceedances are symmetric at a fixed point of time, but they are varying in time through the self-exciting component.

The model proposed in this paper can be applied in various fields, where the analysis of rare events is required. In portfolio models, for instance, where combining

several risky securities it is of great importance to estimate not only the marginal default probabilities, but also (even much more importantly) the joint default probabilities. The model can also be applied as a stress testing tool. Both through simulation and through evaluating different stress scenarios our model yields an approach how banks or insurance companies may prepare their portfolios against extreme events. In general, the model can potentially be applied well beyond the financial context.

References

Ait-Sahalia, Y., J. Cacho-Diaz, and R. J. Laeven (2011): "Modeling Financial Contagion Using Mutually Exciting Jump Processes," *NBER Working Paper No. w15850*.

Bowsher, C. G. (2007): "Modelling security market events in continuous time: Intensity based, multivariate point process models," *Journal of Econometrics*, 141, 876–912.

Chavez-Demoulin, V. (2005): "Estimating Value-at-Risk: A point process approach," *Quantitative Finance*, 5(2), 227–234.

Chavez-Demoulin, V., and P. Embrechts (2011): "An EVT primer for credit risk," *The Oxford Handbook of Credit Derivatives*, 73(1), 500–532.

Coles, S. G. (2001): *An Introduction to Statistical Modeling of Extreme Values*. Springer.

Coles, S. G., and J. A. Tawn (1991): "Modelling extreme multivariate events," *Journal of Royal Statistical Society B*, 53(2), 377–392.

de Haan, L., and A. Ferreira (2006): *Extreme Value Theory: An Introduction*. New York: Springer.

Demarta, S., and A. J. McNeil (2005): "The t copula and related copulas," *International Statistical Review*, 73(1), 111–129.

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

Embrechts, P., T. Liniger, and L. Lin (2011): "Multivariate Hawkes Processes: an Application to Financial Data," *Journal of Applied Probability*, 48(A), 367–378.

Errais, E., K. Giesecke, and L. R. Goldberg (2010): "Affine Point Processes and Portfolio Credit Risk," *SIAM Journal of Financial Mathematics*, 1, 642–665.

Hawkes, A. G. (1971): "Point Spectra of Some Mutually Exciting Point Processes," *Journal of the Royal Statistical Society B*, 33(3), 438–443.

Leadbetter, M. R. (1991): "On a basis for "Peaks over Threshold" modeling," *Statistics and Probability Letters*, 12, 357–362.

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

Resnick, S. I. (2007): *Extreme Values, Regular Variation, and Point Processes*. Springer.

Resnick, S. I., and C. Starica (1995): "Consistency of Hill's estimator for dependent data," *Journal of Applied Probability*, 32.

Rootzen, H., and N. Tajvidi (2006): "Multivariate generalized Pareto distributions," *Bernoulli*, 12, 917–930.

Smith, R. L., J. A. Tawn, and S. G. Coles (1997): "Markov chain models for threshold exceedances," *Biometrika*, 84(2), 249–268.

Tawn, J. A. (1990): "Modelling multivariate extreme value distributions," *Biometrika*, 77,2, 245–253.