

Bayesian smoothing for time-varying extremal dependence

Junho Lee

Financial Supervisory Service, South Korea

Miguel de Carvalho

*School of Mathematics, University of Edinburgh, Edinburgh, UK
CIDMA, Universidade de Aveiro, Portugal*

António Rua

*Banco de Portugal, Economic Research Department, Lisbon, Portugal
Nova School of Business and Economics, Lisbon, Portugal*

Julio Avila

Pontificia Universidad Católica de Chile, Santiago, Chile

Summary. We propose a Bayesian time-varying model that learns about the dynamics governing joint extreme values over time. Our model relies on dual measures of time-varying extremal dependence, that are modelled via a suitable class of generalized linear models conditional on a large threshold. The simulation study indicates that the proposed methods perform well in a variety of scenarios. The application of the proposed methods to some of the world's most important stock markets reveals complex patterns of extremal dependence over the last 30 years, including passages from asymptotic dependence to asymptotic independence.

Keywords: Asymptotic (in)dependence, Bayesian P-splines, Coefficient of tail dependence, Extreme value theory, International equity markets, Nonstationary extremes, Risk management, Statistics of multivariate extremes, Time-varying extremal dependence.

1. Introduction

1.1. Motivation and scope

The collapse of global financial markets has been in full swing due to the global COVID-19 pandemic, rising inflation, and a supply chain shortage. Motivated by the need to understand the comovements of extreme losses in international financial markets over such periods of turbulence, the last few decades have seen a large increase in statistical methods for multivariate extreme values applied to financial markets (e.g. Longin and Solnik, 2001; Embrechts et al., 2002; Poon et al., 2003, 2004; Ergen, 2014; Castro et al., 2018; Gong and Huser, 2022). Key goals of statistical methods for multivariate extreme values include: i) learning about the dependence between the extreme values of a random vector; and ii) extrapolating into the joint tail beyond observed data (Coles 2001, Ch. 8; Beirlant et al. 2004, Ch. 9). Such methodologies for multivariate extreme values are pertinent for virtually any field where there is an interest in quantifying the frequency and magnitude of extreme events and natural hazards, such as in climatology, geology, and hydrology, among others.

Address for correspondence: M. de Carvalho, James Clerk Maxwell Building, The King's Buildings Peter Guthrie Tait Road, Edinburgh, EH9 3FD UK.

E-mail: Miguel.deCarvalho@ed.ac.uk

1.2. Preparations and starting point

In this paper we propose a Bayesian time-varying model that learns about the dynamics governing joint extreme values over time. Our starting point for modelling will be a pair of time-varying versions (see Section 2.1 below) of mainstream measures of tail dependence on which we provide some context next. Suppose that X and Y have common marginal distributions and that common upper endpoint is infinity. Then, the tail dependence measures χ and $\bar{\chi}$ are respectively defined as

$$\chi = \lim_{u \rightarrow \infty} P(X > u \mid Y > u), \quad \bar{\chi} = \lim_{u \rightarrow \infty} \frac{2 \log P(X > u)}{\log P(X > u, Y > u)} - 1. \quad (1)$$

For further details see Coles et al. (1999). If $\chi = 0$, then X and Y are said to be asymptotically independent, whereas if $\chi \in (0, 1]$, they are asymptotically dependent. The coefficient $\bar{\chi} \in (-1, 1]$ quantifies the degree of dependence within the class of asymptotically independent random variables; positive and negative extremal association, and near independence, respectively correspond to $\bar{\chi} \in (0, 1)$, $\bar{\chi} \in (-1, 0)$, and $\bar{\chi} = 0$, whereas $\bar{\chi} = 1$ implies asymptotic dependence (Coles et al., 1999). For example, for bivariate Normal random variables (X, Y) with a positive correlation $\rho \in (0, 1)$, the extremal coefficients are computed as $\chi = 0$ and $\bar{\chi} = \rho$ respectively, which means X and Y are asymptotically independent and their extreme values are positively associated. When studying extremal dependence, it is common to standardise the margins, and here we set the margins to be unit Fréchet, that is, $P(X > u) = P(Y > u) = \exp(-1/u)$, for $u > 0$. Finally, beyond χ and $\bar{\chi}$ our developments will also take advantage of a third measure of extremal dependence, known as the coefficient of tail dependence, $\eta \in (0, 1]$, which follows from the specification (Ledford and Tawn, 1996):

$$P(X > u, Y > u) = \frac{L(u)}{u^{1/\eta}}, \quad u > 0. \quad (2)$$

Here, $L(u)$ is a slowly varying function ($L(uz)/L(u) \rightarrow 1$ for any $z > 0$, as $u \rightarrow \infty$), and X and Y are unit Fréchet-distributed. It can be easily shown that (1) and (2) yield:

$$\chi = \lim_{u \rightarrow \infty} u^{1-1/\eta} L(u), \quad \bar{\chi} = 2\eta - 1. \quad (3)$$

Thus, if $\eta = 1$ and $L(u) \rightarrow c \in (0, 1]$, then $\bar{\chi} = 1$ and X and Y are asymptotically dependent with level $\chi = c$. If, however, $0 < \eta < 1$ (or $\eta = 1$ and $L(u) \rightarrow 0$), then $\chi = 0$ and X and Y are asymptotically independent with level $\bar{\chi} = 2\eta - 1$.

1.3. Main contributions

Our paper contributes to the recent literature on models for multivariate nonstationary extremes that can be used for tracking the dynamics governing extremal dependence over time as well as for assessing the impact of covariates on the extremal dependence structure (e.g., de Carvalho, 2016; Mhalla et al., 2017; Castro et al., 2018; Escobar-Bach et al., 2018; Mhalla et al., 2019a,b). These methods attempt to track changes over time in the extremal dependence structure in an asymptotic dependence context. Since most available methods for nonstationary joint extremes are grounded on multivariate extreme value (MEV) distributions, they are fundamentally tied to an asymptotic dependence framework. Yet, when the series of interest are asymptotically independent, MEV-based approaches are known to yield an over-estimation of the probability of joint extremes, and thus of financial risk (Poon et al., 2003). To our knowledge, the only available method for multivariate nonstationary extremes that deals with asymptotic dependence and asymptotic independence is the recent paper by Mhalla et al. (2019b) that proposes a sturdy exceedance-based nonlinear regression model for tail dependence.

Here, we devise methods that track the dynamics of extremal dependence by adopting smooth time-varying indicators that learn about the dynamics of tail dependence between two stochastic processes. A brief outline of the proposed approach is as follows. We employ time-varying versions of the dual coefficients of extremal dependence in (1), so to devise a Bayesian time-varying model that learns about the dynamics of extremal dependence over time. For modelling, we propose a semi-parametric approach for both time-varying dual coefficients of extremal dependence through a composition of two functions, $F\{g(t)\}$, where F is an inverse link function, and g a smooth function modelled by a family of basis functions. To model the nonstationary patterns of the time-varying dual measures of extremal dependence we employ Bayesian P-splines (Lang and Brezger 2004, Fahrmeir et al. 2011, Section 2.2); Bayesian P-splines are a mainstream Bayesian smoothing methodology with applications in a wealth of contexts (e.g. Bruno et al., 2016; Rondon and Bolfarine, 2016; Sriram et al., 2016; Eilers and Marx, 2021).

Contrarily to most existing approaches for nonstationary joint extremes, the proposed methods are naturally tailored for both asymptotic dependence and asymptotic independence. In addition, since our methods are not based on the MEV distribution, they do not lead to an over-estimation of probability of joint extremes under a framework of asymptotic independence. While the herein proposed methods are as easy to fit as a GLM (generalized linear model) (McCullagh and Nelder, 1989; Dobson, 2008), they will be shown to have a comparable, if not superior, performance in comparison with those of Mhalla et al. (2019b) when the task of interest is to learn about time-varying versions of χ and $\bar{\chi}$, that are formally introduced in Section 2.1. Another merit of the proposed methods is that they directly define a prior on the space of our parameters of interest (i.e., time-varying versions of χ and $\bar{\chi}$). A Bayesian version of Mhalla et al. (2019b) for the same task would entail defining a flexible prior on the space of all Pickands functions, which is a nontrivial task as can be seen from Marcon et al. (2016).

The outline of this article is as follows. Section 2 introduces the proposed Bayesian smoothing methodologies for time-varying extremal dependence. Section 3 presents simulation studies to assess the performance of our methods. In Section 4 we examine an application in international stock markets. Final remarks are given in Section 5.

2. Bayesian smoothing for time-varying extremal dependence

2.1. Dual measures of time-varying extremal dependence

In this section we devise time-varying versions of the dual measures of extremal dependence in Eq. (1). Before introducing the proposed concepts we first lay the groundwork. Observations will be assumed to arise from a discrete-time bivariate stochastic process $\{(X_t, Y_t)\}_{t=1}^n$ with standard unit Fréchet marginal distributions, that is, $X_t \sim F_{X_t}$ and $Y_t \sim F_{Y_t}$ with $F_{X_t}(x) = F_{Y_t}(x) = \exp(-1/x)$, for $x > 0$ and $t \in \{1, \dots, n\}$. The process $\{(X_t, Y_t)\}_{t=1}^n$ is not necessarily stationary in its dependence structure, and thus the tail dependence measures in Eq. (1) are not constant and may depend on time. Motivated by this, we define the time-varying coefficient of extremal dependence χ_t as a function on $[1, n]$ such that

$$\chi_t = \lim_{u \rightarrow \infty} P(X_t > u \mid Y_t > u), \quad (4)$$

for $1 \leq t \leq n$. The interpretation of χ_t is tantamount to that of χ in Eq. (1), but it accounts for the evolution of extremal dependence over time. The measure χ_t can be regarded as a natural extension of the unconditional coefficient χ to a time-changing setup and it is tailored for assessing the extremal dependence at time t after the margins of $\{X_t\}$ and $\{Y_t\}$ have been converted to a common scale. Similarly, a time-varying extremal dependence coefficient

$\bar{\chi}_t$ can be readily defined as

$$\bar{\chi}_t = \lim_{u \rightarrow \infty} \frac{2 \log P(X_t > u)}{\log P(X_t > u, Y_t > u)} - 1, \quad (5)$$

for $1 \leq t \leq n$. In practice, it is often convenient to convert time to the unit interval, in which case both χ_t and $\bar{\chi}_t$ become functions on the unit interval.

Combining the information from the coefficients χ_t and $\bar{\chi}_t$ allows for examining the dependence structure over time in terms of asymptotic dependence and asymptotic independence. For instance, if χ_t grows over a certain period, while $\bar{\chi}_t$ remains at one, which suggests that the extremal dependence between the random processes $\{X_t\}$ and $\{Y_t\}$ has strengthened over that period. On the other hand, if $\{X_t\}$ and $\{Y_t\}$ are asymptotically independent for all t , so that $\chi_t = 0$, then an increasing trend in $\bar{\chi}_t$ should be understood as an increase in extremal dependence.

For illustrating the time-varying characteristics of extremal dependence, we consider an example based on a well-known multivariate extreme value distribution.

EXAMPLE 1 (LOGISTIC FAMILY). The time-varying bivariate extreme value distribution with logistic dependence structure is given by

$$G_t(x, y) = \exp\{-(x^{-1/\alpha_t} + y^{-1/\alpha_t})^{\alpha_t}\}, \quad x > 0, y > 0,$$

where perfect independence between X_t and Y_t arises if $\alpha_t = 1$, otherwise they are asymptotically dependent for $0 < \alpha_t < 1$, with $0 \leq t \leq 1$. It can be shown that

$$\chi_t = 2 - 2^{\alpha_t}, \quad \bar{\chi}_t = 1.$$

Fig. 1 illustrates an example of a time-varying χ_t where the dependence parameter α_t changes between 0.25 and 0.75, i.e. $\alpha_t = 1/4 \sin(2\pi t) + 1/2$, while $\bar{\chi}_t = 1$ across the same period, which indicates asymptotic dependence between X_t and Y_t with the pattern of extremal dependence being cyclical as depicted in the chart of χ_t .

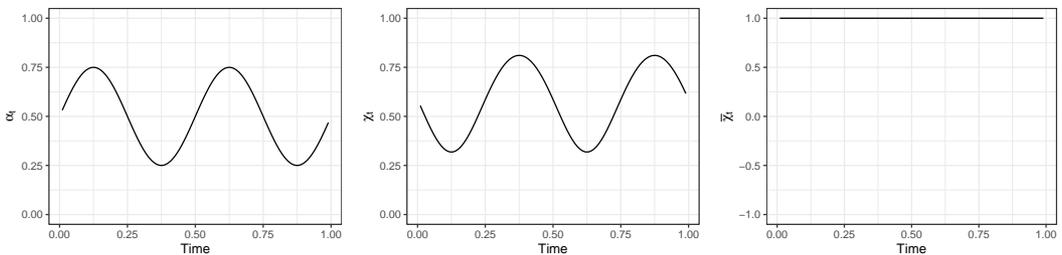


Fig. 1. The evolution of α_t , χ_t , and $\bar{\chi}_t$ over time in the case of asymptotic dependence from the bivariate extreme value distribution with the time-varying logistic dependence structure from Example 1.

2.2. Modelling time-varying coefficients of extremal dependence

This section deals with modelling the extremal coefficients of dependence in Eqs. (4) and (5).

Specification for χ_t

Our specification for χ_t will be semiparametric and it will entail an inverse link function, $F : \mathbb{R} \rightarrow [0, 1]$, and a smooth function $g(t) := g(t; \theta_g)$ that maps the real line into itself,

and which is parameterised by a family of continuous basis functions with a parameter space Θ . Specifically, we set

$$\chi_t \approx P(X_t > u \mid Y_t > u) \equiv \phi_{\chi_t} = F\{g(t)\}, \quad (6)$$

as $u \rightarrow \infty$. The inverse link function F enforces the parametric constraint that the conditional survival probability is contained between 0 and 1, i.e. $0 \leq \chi_t \leq 1$, whereas the smooth function g reflects the effect of time on the tail dependence.

Specification for $\bar{\chi}_t$

For setting up a model for $\bar{\chi}_t$, we do not apply the definition in Eq. (5) directly, but instead employ the so-called coefficient of tail dependence $\eta \in (0, 1]$ (Ledford and Tawn, 1996), and draw $\bar{\chi}_t$ based on the property that $\bar{\chi} = 2\eta - 1$. Let $Z_t = \min\{X_t, Y_t\}$. We specify,

$$P(X_t > u, Y_t > u) = P(Z_t > u) = L_t(u)u^{-1/\eta_t}, \quad (7)$$

where $L_t(u)$ is a time-changing slowly varying function, i.e. $\lim_{u \rightarrow \infty} L_t(uz)/L_t(u) = 1$ for any $z > 0$ and all $t \in [0, 1]$. The specification in (7) can be seen as an extension of that of Ledford and Tawn (1996, 1997) to a time-varying setting, or as an extension of the univariate tail index regression approach Wang and Tsai (2009) to a framework of multivariate extremes. We model a time-varying version of the tail dependence coefficient η_t using a similar construction as that for ϕ_{χ_t} . That is, let $\phi_{\eta_t} = E\{\log(Z_t/u) \mid Z_t > u\}$, i.e., ϕ_{η_t} is a time-varying mean excess function, from which the Hill estimator is often derived (Beirlant et al., 2004, p. 104), and that converges to η_t in Eq. (7), as $u \rightarrow \infty$. Then, we set

$$\eta_t \approx \phi_{\eta_t} = H\{l(t)\}, \quad (8)$$

for all t where $H : \mathbb{R} \rightarrow [0, 1]$ is an inverse link function, and $l(t) := l(t, \boldsymbol{\theta}_t)$ is a smooth function, that maps the real line into itself, and which is parameterised by a family of basis functions with a parameter space Θ . Likewise as above, the link H enforces the parametric constraint, $0 < \eta_t \leq 1$, whereas the smooth function l controls the dynamics of extremal dependence over time. Our specification for $\bar{\chi}_t$ follows directly from (8) and it is given by

$$\bar{\chi}_t \approx 2H\{l(t)\} - 1, \quad (9)$$

as $u \rightarrow \infty$. In line with Poon et al. (2003, p. 938) we recommend to focus on χ_t only if there is evidence in favor of $\bar{\chi}_t = 1$; otherwise, the focus should be placed on $\bar{\chi}_t$.

After introducing specifications (6) and (9), we are now ready to discuss how to learn about these measures of extremal dependence from data.

2.3. Observation models and completing model specification

The first part of this section derives observation models for ϕ_{χ_t} and ϕ_{η_t} , whereas the second part comments on how to complete the model specification. We start with ϕ_{χ_t} . Let $\{I_t\} = \{\mathbb{1}_{\{X_t > u\}} : Y_t > u\}$, where $\mathbb{1}_A$ is the indicator function. Then, I_t is approximately Bernoulli distributed with mean ϕ_{χ_t} , for a large u . To obtain the sampling distribution about ϕ_{η_t} , let $\{E_t\} = \{\log(Z_t/u) : Z_t > u\}$. Then it can be shown that E_t given $Z_t > u$ follows an Exponential distribution with mean η_t , as $u \rightarrow \infty$. The observation models for ϕ_{χ_t} and ϕ_{η_t} are thus summarised as

$$I_t \sim \text{Bern}(\phi_{\chi_t}), \quad E_t \sim \text{Exp}(\phi_{\eta_t}), \quad (10)$$

for all t . The inference goal is to learn about χ_t and $\bar{\chi}_t$ as specified in Eqs. (6) and (9) from $k_I = |\{I_t\}|$ and $k_E = |\{E_t\}|$ pseudo-observations from $\{I_t\}$ and $\{E_t\}$ (respectively), with $|\cdot|$ denoting cardinality.

To complete the model specification some comments on the link functions and on the smooth functions are in order. Similarly to the setup of generalized linear models (GLMs) (McCullagh and Nelder, 1989; Dobson, 2008) the link functions F and H are set in advance by the user. We then learn about the smooth functions g and l through the connection between the sampling distribution of observation models and the specifications in Eq. (6) and Eq. (8). Since the Bernoulli and Exponential distributions are both members of the Exponential family, estimation of g and l can be framed into a GLM setting.

We complete the model specification by modelling the basis functions using B-splines (de Boor, 2001). Consider $m + 1$ equally-spaced knots, $t_0 < \dots < t_m$. The smooth functions are then modelled as

$$g(t) = \sum_{k=1}^K \beta_k^{(g)} B_k^d(t), \quad l(t) = \sum_{k=1}^K \beta_k^{(l)} B_k^d(t), \quad (11)$$

where $B_k^d(t)$ is a B-spline basis function of degree d evaluated at time t and $K = d + m$.

We next discuss details on the prior specification for $\boldsymbol{\beta}^{(g)} = (\beta_1^{(g)}, \dots, \beta_K^{(g)})^\top$ and $\boldsymbol{\beta}^{(l)} = (\beta_1^{(l)}, \dots, \beta_K^{(l)})^\top$ in $\Theta = \mathbb{R}^K$ as well as on posterior inference.

2.4. Prior specification and posterior inference

To learn about the coefficients of the B-splines in (11) we use the Bayesian P-spline approach of Lang and Brezger (2004). The Bayesian penalised version of B-splines controls the roughness of a fitted curve by penalising differences of adjacent B-spline coefficients; this is achieved by putting a random walk prior on the first or the second order difference between the coefficients in Eq. (11). To ease notation we focus on presenting the details for a single $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)^\top$, rather than for both $\boldsymbol{\beta}^{(g)}$ and $\boldsymbol{\beta}^{(l)}$. We assign a first-order random walk prior to the coefficient vector $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)^\top$ of each smooth function, which specifies a priori that the neighbouring components of $\boldsymbol{\beta}$ are related via an independent and identical Gaussian error ε_k with mean zero and variance τ^2 ; that is, we set

$$\beta_k = \beta_{k-1} + \varepsilon_k, \quad \varepsilon_k \sim N(0, \tau^2), \quad k = 2, \dots, K, \quad (12)$$

with a flat (uniform) prior for the initial coefficient β_1 . The first order random walk prior can be represented in a matrix form $\mathbf{D}\boldsymbol{\beta} = \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} = (\varepsilon_2, \dots, \varepsilon_K)^\top$, and \mathbf{D} is a difference matrix of dimension $(K - 1) \times K$. The matrix \mathbf{D} has 1's in diagonal elements ($i = j$), -1 's in the next elements from the diagonal ($i = j + 1$), and zeros otherwise, for $i = 1, \dots, K - 1$, and $j = 1, \dots, K$. The variance τ^2 controls the degree of smoothness of the smooth function (say, g or l); a small value of τ^2 results in a less wiggly curve, as each component of $\boldsymbol{\beta}$ tends to be close to the value of its neighbouring component. Accordingly, the conditional probability of the regression coefficients $\boldsymbol{\beta}$ given τ^2 is

$$\pi(\boldsymbol{\beta} \mid \tau^2) \propto \exp\left(-\frac{1}{2\tau^2} \boldsymbol{\beta}^\top \mathbf{K} \boldsymbol{\beta}\right), \quad (13)$$

where \mathbf{K} is a penalty matrix, $\mathbf{K} = \mathbf{D}^\top \mathbf{D}$. In the full Bayesian setting, the precision parameter τ^2 is also estimated along with the regression coefficients by assigning an hyper-prior distribution to it. We place a diffuse Inverse Gamma prior $\tau^2 \sim \text{IG}(a_0, b_0)$ with hyper-parameters $a_0 > 0$ and $b_0 > 0$. The joint posterior distribution is

$$\mathbf{p}(\boldsymbol{\beta}, \tau^2 \mid \{I_t\}, \{E_t\}) \propto L(\boldsymbol{\beta} \mid \{I_t\}, \{E_t\}) \pi(\boldsymbol{\beta} \mid \tau^2) \pi(\tau^2).$$

Bayesian inference can then be based on

$$\mathbf{p}(\boldsymbol{\beta} \mid \{I_t\}, \{E_t\}) \propto L(\boldsymbol{\beta} \mid \{I_t\}, \{E_t\}) \pi(\boldsymbol{\beta} \mid \boldsymbol{\tau}^2), \quad \mathbf{p}(\boldsymbol{\tau}^2 \mid \boldsymbol{\beta}) \propto \pi(\boldsymbol{\beta} \mid \boldsymbol{\tau}^2) \pi(\boldsymbol{\tau}^2).$$

The full conditional distribution $\mathbf{p}(\boldsymbol{\tau}^2 \mid \boldsymbol{\beta})$ is a conjugate Inverse Gamma posterior. The posterior of the coefficients of the smooth functions, $\mathbf{p}(\boldsymbol{\beta} \mid \{I_t\}, \{E_t\})$ is not however analytically tractable, and thus we employ Markov Chain Monte Carlo (MCMC) methods to sample from it. Posterior sampling can be conducted via a Metropolis–Hastings algorithm with IWLS (Iteratively Weighted Least Squares) proposals (Brezger and Lang, 2006). Further details are available in the Supporting Information.

We close this section with some comments on standard model choices in terms of the number of knots ($m + 1$) as well as for the degree of the spline function. The basic principle underlying P-splines is to use a large number of inner knots (typically 20–40) and then to penalize so to avoid overfitting (Fahrmeir et al., 2011, p. 43). While approaches for optimizing the number of knots have been devised (Ruppert, 2002) their added value in practice is not clear, and it has been heavily criticized by others (e.g., Eilers and Marx, 2021, p. 160). We now move to the order of the spline. Setting, for example, $d = 0$ would result in fits that are piecewise constant and non-differentiable functions with the same number of steps as the number of inner knots. Setting $d \geq 1$ leads to fits that are continuous and $d - 1$ differentiable functions. It is often argued (e.g., Fahrmeir et al., 2011, p. 41) that for many applications cubic polynomial splines ($d = 3$) are an appropriate choice, one that leads to a continuous and twice differentiable function.

3. Simulation study

3.1. Description of scenarios and preliminary numerical experiments

We conduct a simulation study to assess the performance of the proposed methods. We start by describing the data-generating processes and then show results from a one-shot experiment; Monte Carlo evidence will be reported in Section 3.2. We examine the performance of the proposed methods over four simulation scenarios described in Table 1. For Scenarios A–D in Table 1, we assume that the time-varying parameter controlling the degree of extremal dependence is

$$\theta(t) = \{1 - (1 - t)^2\} [\cos\{4\pi(1 - t)^2\} + 1]/4 + 0.2, \quad 0 \leq t \leq 1.$$

We draw a random sample from the joint distribution for each scenario of size $n = 40\,000$, then transform their marginal distributions into standard unit Fréchet distributions, and finally threshold the data at 0.95 quantiles of Y_t (for $\{I_t\}$) and of Z_t (for $\{E_t\}$). As shown in the Supporting Information, the performance of the proposed methods is also overall satisfactory for a sample of size 10 000 with a 0.95 threshold. To complete the specification of our model, we assign a diffuse Gamma distribution for the hyper-prior of $\boldsymbol{\tau}^2$ and use B-spline basis functions with $m + 1 = 20$ knots and degree $d = 3$ (cubic splines). Posterior samples are gathered by means of MCMC methods, with a burn-in period of 10 000, where we save each 10 iterations with two chains, thus obtaining an MCMC sample of size 2 000 for all the parameters and function evaluations of interest.

Fig. 2 depicts the results of a single-run experiment for each scenario; each panel contains the scatterplot of the log transformed original data and the posterior mean χ_t and $\bar{\chi}_t$. As it can be seen from Fig. 2 all posterior mean fits recover reasonably well the true targets. Despite the overall accuracy of the estimates, the fit of χ_t for Scenario D is not as good as that for the remaining scenarios—although it accurately recovers the sub-asymptotic $\chi_t(u) \equiv P\{X_t > u \mid Y_t > u\}$. As mentioned by a reviewer, the challenge is that whenever a

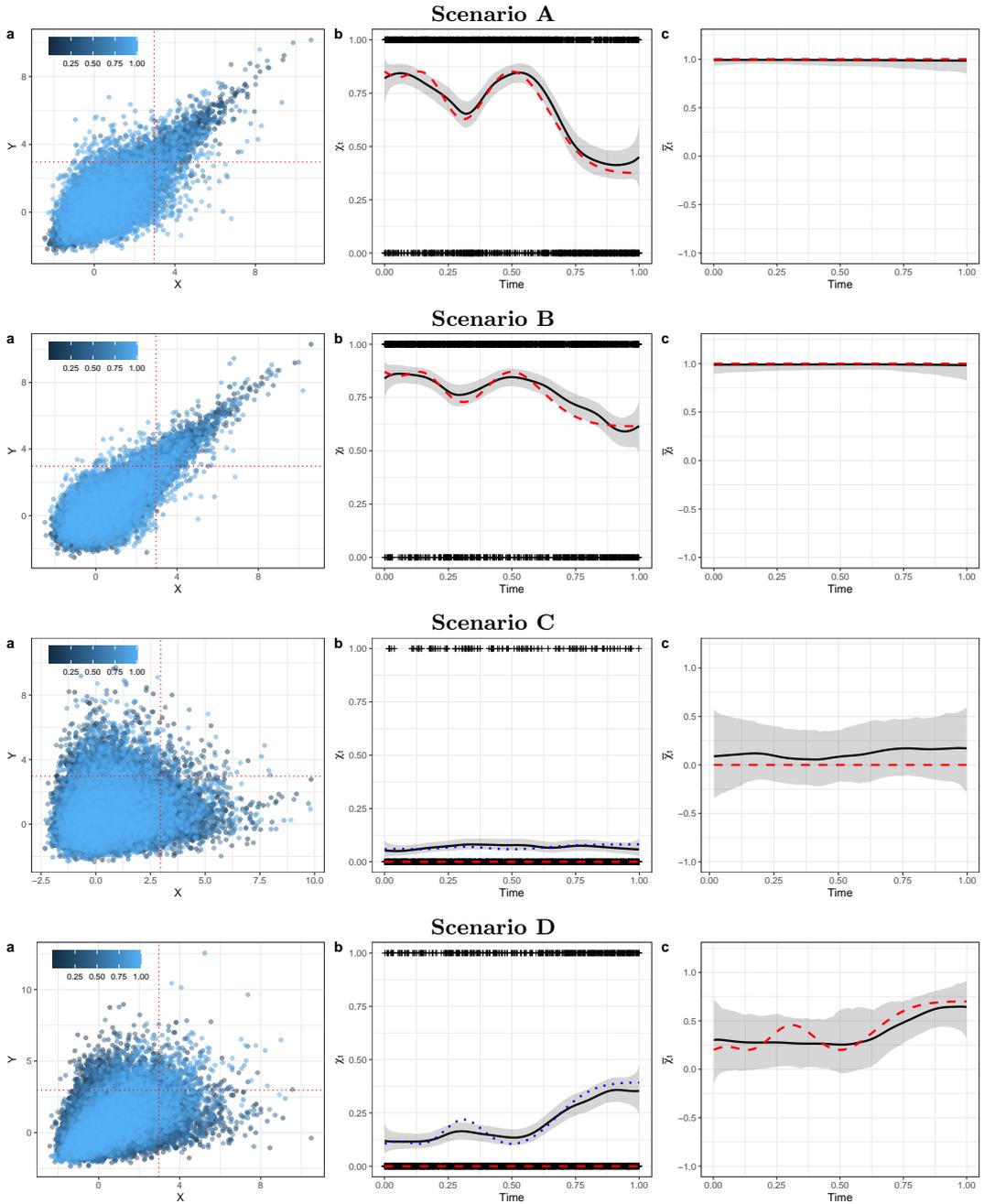


Fig. 2. Single-run numerical experiment. Left: Scatterplot of log transformed data and posterior distribution of time-varying χ_t and $\bar{\chi}_t$ for Scenarios A–D. Middle and right: Posterior mean χ_t and $\bar{\chi}_t$ (solid) and credible bands are presented along with the true (dashed). The true sub-asymptotic $\chi_t(u)$ (dotted) is also represented for Scenarios C–D. The rug in the middle panel corresponds to the points $\{(t, \mathbb{1}_{\{\chi_t > u\}}) : Y_t > u\}$.

Table 1. Simulation scenarios. AI and AD respectively denote asymptotic independence and asymptotic dependence, and the joint distribution functions are defined on the unit square, i.e., $(u, v) \in [0, 1]^2$.

Scenario	Copula	Distribution function	Parameter	Class
A	Logistic family	$\exp[-\{(-\log u)^{\delta_t} + (-\log v)^{\delta_t}\}^{1/\delta_t}]$	$\delta_t = \theta(t)$	AD
B	Clayton lower tail	$(u^{-\delta_t} + v^{-\delta_t} - 1)^{-1/\delta_t}$	$\delta_t = \theta(t)$	AD
C	Morgenstern	$uv\{1 + \alpha_t(1-u)(1-v)\}$	$\alpha_t = \theta(t)$	AI
D	Bivariate Normal	$\Phi_2\{\Phi^{-1}(u), \Phi^{-1}(v); \rho_t\}$	$\rho_t = \theta(t)$	AI

distribution has ‘strong’ asymptotic independence, represented by moderate-to-large values of η_t , then the sub-asymptotic $\chi_t(u)$ is typically very far from its limiting value of zero. An additional numerical experiment in the Supporting Information confirms this is indeed the case.

Thus, exceptionally in Fig. 2, and for Scenario D, the posterior mean fits for χ_t are also compared with the true χ_t at the sub-asymptotic level at which we threshold the data; that is, we also compare the fitted χ_t for Scenario D with

$$\chi_t(u) = 2 - \frac{\log C_t(u, u)}{\log u},$$

where we fix $u = 0.95$. Here, $C_t(u, v) = \Phi_2\{\Phi^{-1}(u), \Phi^{-1}(v); \rho_t\}$ is the time-varying copula associated to the bivariate Normal distribution, for $(u, v) \in (0, 1)^2$, and Φ_2 and Φ are the distribution functions of the unit variance bivariate Normal and the standard Normal, respectively. As it can be noticed from Fig. 2, for Scenario D the posterior mean χ_t is not far from the true target, $\chi_t = 0$, but it is much closer to its sub-asymptotic version $\chi_t(u)$.

To sum up, overall the single-run experiment anticipates an overall satisfactory fit of the proposed methods, and it also foresees that fitting χ_t in Scenario D is more demanding due to the log-convergence rate of the slowly-varying function. Of course such findings should be regarded as tentative at this stage, as this is just the outcome of a single run experiment, and the goal of the next section will be to inspect this further.

3.2. Monte Carlo evidence

A Monte Carlo simulation study was conducted to evaluate the performance of our methods, in a battery of experiments that extend the numerical analysis from Section 3.1. Specifically, we repeat the one shot experiments from Section 3 $R = 500$ times. The Monte Carlo mean of the posterior means reported in the Supporting Information, reveals an overall satisfactory performance of the proposed methods at recovering the true targets. In common with all extreme value theory methodologies there is however a price to pay due to the fact that we are not yet in the limiting case ($u \rightarrow \infty$).

We have also conducted a battery of additional numerical experiments to compare the proposed methods with exceedance-based regression methods of Mhalla et al. (2019b). Specifically, we estimated the Pickands dependence function and the angular dependence function

Table 2. MIAE (Mean Integrated Absolute Error) from the numerical experiments between the proposed methods and Mhalla et al. (2019b).

Scenario	Proposed methods				Mhalla et al. (2019b)			
	A	B	C	D	A	B	C	D
χ_t	0.0280	0.0215	0.0088	0.0194	0.0761	0.0590	0.0514	0.1557
$\bar{\chi}_t$	0.0108	0.0079	0.1016	0.1125	0.4680	0.8866	0.0454	0.1002

from the exceedance-based regression methods of Mhalla et al., and then used those functions to produce plug-in estimators for the time-varying dual measures χ_t and $\bar{\chi}_t$. See the Supporting Information as well as Table 2 where we report the posterior MIAE (Mean Integrated Absolute Error), defined by

$$\text{MIAE}(\chi) = E\left(\int_0^1 |\chi_t - \chi_t^*| dt \mid \{I_t\}\right), \quad \text{MIAE}(\bar{\chi}) = E\left(\int_0^1 |\bar{\chi}_t - \bar{\chi}_t^*| dt \mid \{E_t\}\right).$$

Here, χ_t^* and $\bar{\chi}_t^*$ are the true coefficients values. Lower values of the MIAE, indicate a better performance. Hence, as can be seen from Table 2 the proposed methods have a comparable, if not superior, performance in comparison with those of Mhalla et al. (2019b). Beyond the battery of numerical experiments reported above, we also document in the Supporting Information evidence suggesting a good performance of the proposed methods under time-varying margins.

4. Time-varying extremal dependence in international stock markets

4.1. Financial context and motivation for the empirical enquiry

There have been plenty of studies in the financial literature analysing stock market comovements (e.g., Forbes and Rigobon, 2002; Brooks and Del Negro, 2004; Morana and Beltratti, 2008; Rua and Nunes, 2009; Albuquerque and Vega, 2009; Jach, 2017; Ehrmann and Jansen, 2020). Despite this and the fact that methodologies for modeling dependence over time are available (e.g., Patton, 2006), few attempts have been made to examine the dynamics governing the comovement of extreme values on stock markets over time, and this will be a key goal of our empirical inquiry. An important exception in this regard is the seminal work of Poon et al. (2003, 2004) who provide evidence of increasing levels of extremal dependence among losses on some leading stock markets; yet their subperiod analysis is merely exploratory in the sense that it arbitrarily partitions the sample period into three subperiods. Another exception is Castro et al. (2018) who examine the temporal evolution of extremal dependence of European stock markets using a conditional angular density; since their approach is however based on MEV distributions it is inherently tied to a setup of asymptotic dependence. Below, we use the same settings as in Section 3 in terms of the number of knots and degree ($m + 1 = 20$ and $d = 3$); a sensitivity analysis is reported in the Supporting Information.

A main goal will be to quantify the degree of the market integration in terms of extremal dependence and to unveil the dynamics governing the dependence of extreme losses in global stock markets. From an empirical outlook our analysis extends the seminal papers of Poon et al. in a number of significant ways. Perhaps the most important one is that our analysis tracks the dynamics of extreme value dependence over time, whereas the analysis in Poon et al. is static, although it acknowledges the existence of such dynamics, en passant, over their subperiod analysis. Our analysis will further include China, now a huge player in the global economy, that was not covered in the analysis of Poon et al. Additionally, our inquiry will cover a variety of post 2003–2004 noteworthy events—that have yielded unseen levels

of global financial upheaval—such as the subprime and the European sovereign debt crises, Brexit, and the COVID-19 global pandemic.

We analyse six stock markets across Europe (CAC 40, France; DAX 30, Germany; FTSE 100, UK), East Asia (HANG SENG, China; NIKKEI 225, Japan), and North America (S&P 500, US). The motivation for choosing these stock markets is threefold. First, three of these leading stock markets have been for some time member states of the European Union, and thus are expected to be more integrated; a potential impact of the Brexit referendum will also be under scrutiny in our analysis. Second, the two stock markets from East Asia will allow us to examine how the extremal dependence in European markets compares with that of their East Asian peers. Finally, the evaluation of the extremal dependence of stock markets of the two former regions (i.e., Europe and East Asia) against the US is fundamental, as the US is a key player on the global financial landscape, well-known among other things for having the largest stock markets in the world in terms of equity market value (e.g., Statista, 2021). In line with related studies (e.g., Poon et al., 2003, 2004; Castro et al., 2018), our analysis focuses on pairs of stocks. Note that a pairwise structure completely characterizes the so-called tail-dependence matrix of a d -dimensional vector $\mathbf{X}_t = (X_{1,t}, \dots, X_{d,t})$ (Embrechts et al., 2016, Definition 3.2), and it follows from Berman (1961) that \mathbf{X}_t is asymptotically independent at time t if all pairs $(X_{i,t}, X_{j,t})$ are asymptotically independent, with $i \neq j$. In addition, a pairwise analysis is rather convenient for visualizations. We underscore however that the pairwise structure is insufficient to determine the higher order structure (e.g. not much can be learned about $P(X_{1,t} > u, \dots, X_{d,t} > u)$ from the pairs).

4.2. Data description and summaries for the margins and joint

We retrieve from Datastream closing daily stock index levels for the markets under analysis; the sample period ranges from 5 March 1990 to 4 May 2020. The returns under study obey a variety of well-known empirical properties—such as slightly negative skewness, excess kurtosis, among others—known in financial jargon as stylized facts (e.g., Gentle, 2020, Section 1.6); evidence in favor of this is provided in the Supporting Information (Section 2).

We consider as the unit of analysis negative daily returns of each index, which are defined as the first differences of logarithmic prices, and which can be regarded as a proxy for losses. Following standard practice in related literature, we first filter the data by fitting a GARCH(1,1) model so to remove the heteroskedasticity inherent in each series of stock index returns; for details on GARCH (Generalized Autoregressive Conditional Heteroskedasticity) see, for instance, Koop (2006, Ch. 12). To evaluate how heavy are the tails of the negative log returns in each market, say Y_t , we fit a plain vanilla Pareto-type model with a Jeffreys' prior (e.g. Beirlant et al., 2004, Section 11.5.2), that is,

$$\frac{Y_t}{u} \mid \xi \sim \text{Pareto}(1/\xi), \quad \pi(\xi) \propto \frac{1}{\xi},$$

where u is a large threshold, here considered as the 0.95 quantile of Y_t ; the shape parameter $\xi > 0$ is known in extreme value parlance as the extreme value index, with a higher value indicating an heavier tail. Table 3 presents the estimated posterior mean extreme value indices for all stock markets under study, along with corresponding 95% credible intervals, for both unfiltered and filtered stock index returns. As can be seen from Table 3, the magnitude of extreme losses is comparable across markets, and as expected filtered negative returns present lower extreme value indices.

After transforming the filtered returns of each pair of the stock indices into unit Fréchet margins, we summarise the dual dependence measures over the entire period of analysis in Table 4. The estimates of the coefficients χ and $\bar{\chi}$, as defined in (1), are obtained by their empirical estimators (Beirlant et al., 2004, p. 348) using the 95% quantile. As can be noticed

Table 3. Extreme value indices for (filtered) negative log returns over the entire sample period

	UK	FRA	GER	CHN	JPN	US
Negative log returns						
Posterior mean	0.405	0.373	0.380	0.380	0.357	0.383
Upper 95%	0.370	0.331	0.338	0.352	0.329	0.347
Lower 95%	0.444	0.405	0.411	0.411	0.386	0.422
GARCH-filtered negative log returns						
Posterior mean	0.266	0.266	0.263	0.294	0.300	0.291
Upper 95%	0.242	0.241	0.238	0.272	0.277	0.264
Lower 95%	0.291	0.295	0.289	0.318	0.325	0.320

Table 4. Dependence between negative GARCH-filtered log returns as measured by χ and $\bar{\chi}$

Within-region			Between-regions			With US		
Pair	χ	$\bar{\chi}$	Pair	χ	$\bar{\chi}$	Pair	χ	$\bar{\chi}$
UK–FRA	0.50	0.75	UK–CHN	0.16	0.51	UK–US	0.22	0.63
UK–GER	0.47	0.85	UK–JPN	0.13	0.45	FRA–US	0.21	0.62
FRA–GER	0.54	0.88	FRA–CHN	0.09	0.40	GER–US	0.24	0.72
CHN–JPN	0.23	0.59	FRA–JPN	0.12	0.46	CHN–US	0.08	0.30
			GER–CHN	0.15	0.43	JPN–US	0.10	0.43
			GER–JPN	0.14	0.52			

from Table 4, not surprisingly, both dual measures of extremal dependence tend to be higher on the “within a region” analysis. In addition, for the European markets both measures tend to be higher on the “with US” analysis, in comparison with “between regions”, thus hinting a higher level of extremal dependence between Europe–US in comparison to Europe–East Asia.

Similarly to Poon et al., Table 4 is however static and the next section will extend it to a time-changing context by applying the methods proposed herein.

4.3. Modelling time-varying extremal dependence

In this section we implement the proposed methods to assess how the dependence structure of bivariate extreme losses has been evolving among the six leading stock markets under study over the last three decades. Throughout, we use the same settings for the prior and MCMC as in Section 3. After filtering the negative log returns and converting them to the unit Fréchet scale, we compute a common threshold u applying the 95% quantile to pairwise minima $Z_t = \min\{X_t, Y_t\}$, and obtain the pseudo-samples $\{I_t\}$ and $\{E_t\}$, per market.

In Figs. 3 and 4 we present the dynamics of extremal dependence structure for the pairs of stock indices over the last three decades since 1990. We also compare our results with those of Poon et al. (2003, 2004). Particularly, the values of χ and $\bar{\chi}$ from the subperiod analysis in Poon et al. are overlaid in Figs. 3 and 4 whenever they are available; their subperiod analysis is for each pair of UK, France, Germany, US, and Japan from November 1990 to November 2001. The fits of our measures χ_t and $\bar{\chi}_t$ are roughly of the same magnitude as those obtained earlier by Poon et al. over the subperiod 1990–2001.

To keep track of some noteworthy episodes in the EU agenda over the sample period, as well as their potential for impacting extremal dependence in European markets, we overlay in the charts of χ_t and $\bar{\chi}_t$ the following timelines:

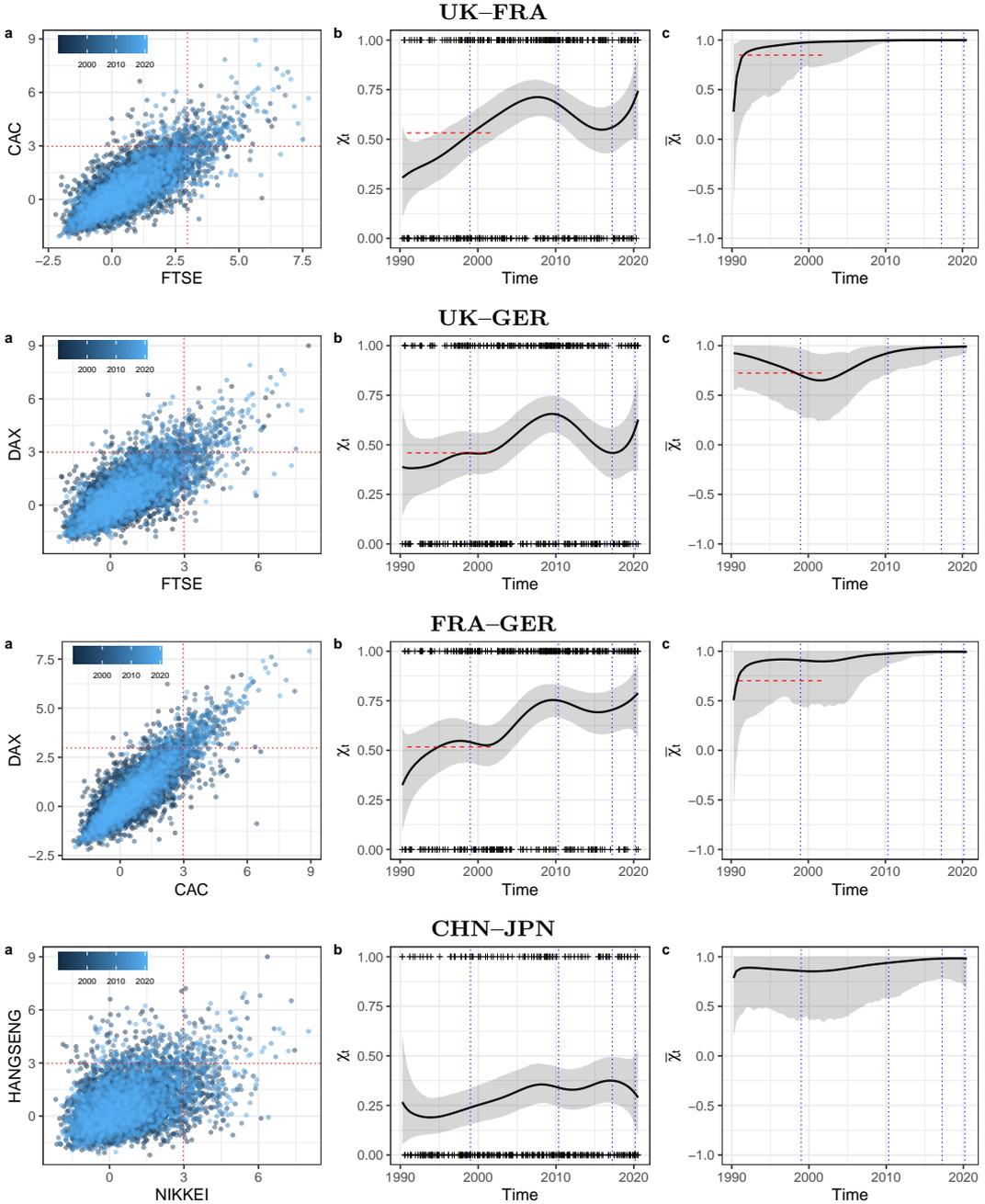


Fig. 3. Within-region analysis. Left: Scatterplot of log transformed data. Middle and Right: Posterior mean estimates for time-varying χ_t and $\bar{\chi}_t$ (solid) along with credible bands; the rug in the middle panel corresponds to the points $\{(t, \mathbb{1}_{\{X_t > u\}}) : Y_t > u\}$ whereas the dashed line corresponds to the available values from the subperiod analysis of Poon et al. (2003). The within-region analysis considers three stocks indices from Europe (CAC, France; DAX, Germany; FTSE, UK) and two from East Asia (HANG SENG, China; NIKKEI, Japan).

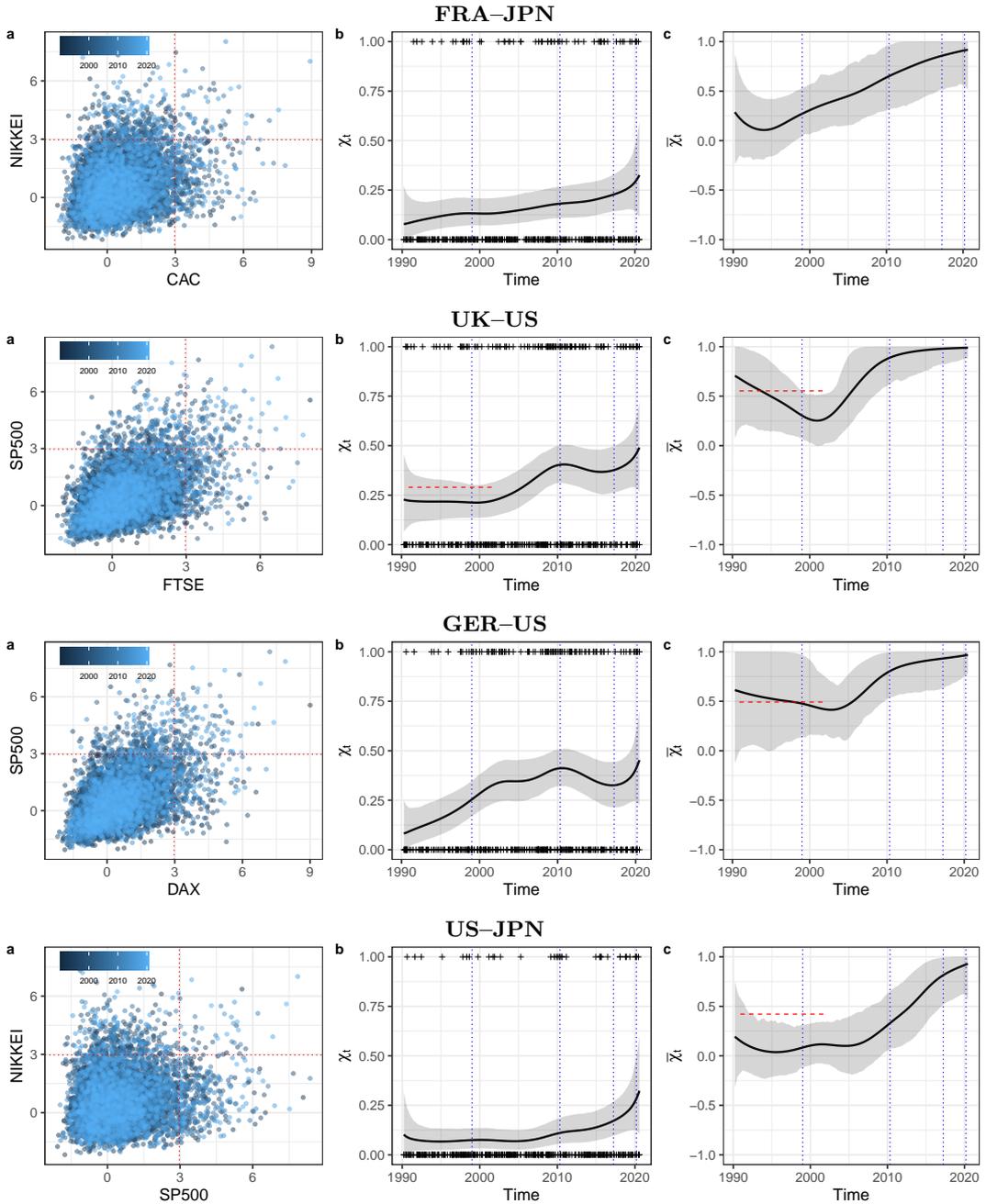


Fig. 4. Between-regions and with US analyses. Left: Scatterplot of log transformed data. Middle and Right: Posterior mean estimates for time-varying χ_t and $\bar{\chi}_t$ (black solid) along with credible bands (grey area); the rug in the middle panel corresponds to the points $\{(t, \mathbb{1}_{\{X_t > u\}}) : Y_t > u\}$ whereas the dashed line corresponds to the available values from the subperiod analysis of Poon et al. (2003).

Reference timelines

- A) Beginning of stage three of European Economic and Monetary Union. 1 January, 1999 (t_A)
- B) Activation of the assistance package for Greece which triggered the European sovereign debt crisis. 2 May, 2010 (t_B)
- C) UK invocation of Article 50 of the Lisbon Treaty for Brexit. 29 March 2017 (t_C)
- D) World Health Organization (WHO)'s confirmation of global pandemic of COVID-19. 11 March 2020 (t_D)

A main goal below is to determine how the extremal dependence of losses on pairs of markets may change over time by comparing periods sufficiently apart from each other. The left panel in Figs. 3 and 4 depicts a scatterplot of the log transformed data as well as the posterior mean χ_t and $\bar{\chi}_t$ fitted using the proposed methods. Some interesting dynamics are unveiled by these charts.

Let's start with the "within-region analysis", that is, with the comparison of markets within Europe and within East Asia. First, as can be seen from Fig. 3, at the beginning of the sample period χ_t presents a similar level of extremal dependence in all pairs of the European stock markets, while their Asian counterparts present moderately lower yet comparable levels; all in all, Fig. 3 suggests thus a mild level of extremal dependence back in the 1990s across all markets. For the first two decades, extreme joint losses of UK–FRA and FRA–GER indices exhibit clear evidence for an increasing level of extremal dependence with the estimated posterior probability of χ_{t_B} to exceed χ_{t_0} being 0.92 and 0.98, respectively. Specifically, the FRA–GER pair manifests stronger extremal dependence levels, thus suggesting a higher degree of association of extreme losses in those markets over recent years. After the European sovereign debt crisis, the level of extremal dependence between the EU states levels off until around the time that Article 50 was triggered by the British (Timeline C). The estimated posterior probability of χ_{t_B} to exceed χ_{t_C} is 0.85 for UK–FRA, 0.70 for UK–GER, and 0.80 for FRA–GER. These estimated posterior probabilities, along with Fig. 3, suggest that after Timeline C there is a moderate increase in extremal dependence in European markets, with less evidence in favor of this for the pair UK–GER. Now, in terms of East Asia, as can be seen in Fig. 4, the degree of association of negative log returns in CHN–JPN has changed much less over time in comparison with their European counterparts. Finally, in the last column of each row of Fig. 3, the coefficient $\bar{\chi}_t$ supplements the extremal dependence measured by the coefficient χ_t with all the coefficient values reasonably close to one, providing thus evidence in favor of asymptotic dependence.

Let's now move to the "between regions" and "with US" analyses; that is, the analysis that compares Europe and East Asia, and finally Europe and East Asia versus the US; keeping in mind space constraints, we only present here part of the analysis and some further charts are reported in the Supporting Information. In Fig. 4, $\bar{\chi}_t$ suggests an increase in extremal dependence across all pairs at least after Timeline A, and it suggests asymptotic dependence over recent years on all pairs. To provide some financial context on what hides behind such rising trend of extremal dependence, let's consider the case of FRA–JPN. As claimed in 2006 by former French President Jacques Chirac (Chirac, 2006):

"Moreover, 45% of the capital of the major French companies in the top 40 listed on the French Stock Exchange (CAC 40) is held by foreign firms—45%, which is a record in Europe."

And Japan is among the main non-resident holders of such capital. Indeed, according to *Banque de France*, Japan has been over 2010–2015 a leading non-resident holder of CAC 40 shares, along with the Euro Area, UK, US, Switzerland, and Canada (Guette-Khiter, 2016, p. 39).

5. Final observations and future research

This paper develops a Bayesian smoothing method that learns from data about the time-changing behaviour of joint extreme values. The proposed model takes full advantage of Bayesian smoothing methods using P-splines to develop flexible yet parsimonious models for a class of dual time-varying extremal dependence measures. Our empirical analysis puts forward how the extremal dependence between joint extreme losses in leading international equity markets has been evolving over the last 30 years in some of the world's most important equity markets. The analysis reveals an increasing trend in the strength of synchronization of crashes on some of these markets as well as some interesting dynamics on the aftermath of the Brexit referendum; in addition, our analysis unveils that some markets may be switching between asymptotic dependence and asymptotic independence. From an empirical perspective our case study expands and updates the influential papers of Poon et al. (2003, 2004) in a number of important ways. Perhaps the most fundamental difference is that our analysis learns about the signature of the dynamics governing extreme value dependence, whereas the analysis in Poon et al. is constant over time. Modelling such dynamics is of chief interest in practice as it reveals how markets move together over black swans, which is key from a risk diversification outlook.

Although we have centered the paper on a financial application, the proposed methods can be readily applied to other fields where the behaviour of extreme events might be expected to change over time (e.g., climatology, geology, hydrology). We underscore that the proposed approach applies to the case where margins are nonstationary provided that the data are preprocessed, for example using splines (e.g. Ferrez et al., 2011, Section 2) or by converting raw nonstationary data into unit Fréchet margins using a time-varying distribution function estimator as exemplified in the Supporting Information. Interestingly, for the inferences, it would seem natural to try to take advantage of the fact that

$$\chi_t = \lim_{u \rightarrow \infty} P(X_t > u \mid Y_t > u) = \lim_{u \rightarrow \infty} P(Y_t > u \mid X_t > u), \quad (14)$$

as $P(X_t > u) = P(Y_t > u) = \exp(-1/u)$, for $u > 0$. While (14) suggests that we could in principle use another sample $\{I'_t\} = \{\mathbb{1}_{\{Y_t > u\}} : X_t > u\}$ for learning about χ_t , since both samples ($\{I_t\}$ and $\{I'_t\}$) are dependent this might impact credible intervals.

Some final words on future research are in order. First, it would be natural to endow the measures χ_t and $\bar{\chi}_t$ with the ability to track structural changes, change points, or other types of abrupt changes in the extremal dependence; despite the fact that breaks in tail behaviour are fundamental in financial applications (e.g., Quintos et al., 2001), most focus has been placed on modeling structural changes on the marginal tail rather than on the joint tail. Second, while here the focus has been on tracking changes on extremal dependence over time, for applied settings where p covariates $\mathbf{x} = (x_1, \dots, x_p)^T$ are available, a generalized additive (Wood, 2017) version of the proposed framework can be readily constructed, as indeed the margins of $\{I_t\}$ and $\{E_t\}$ in (10) are members of the Exponential family; thus, covariate-adjusted coefficients $\chi_{\mathbf{x}}$ and $\bar{\chi}_{\mathbf{x}}$ could more generally be modelled as

$$\chi_{\mathbf{x}} = F\left(\beta_0 + \sum_{j=1}^p \sum_{k=1}^K \beta_k B_k^d(x_j)\right), \quad \bar{\chi}_{\mathbf{x}} = 2H\left(\beta_0 + \sum_{j=1}^p \sum_{k=1}^K \theta_k B_k^d(x_j)\right) - 1. \quad (15)$$

Still within a regression framework, a Bayesian Lasso version of (15) could in principle be devised via the group Lasso (Yuan and Lin, 2006) by shrinking groups of regression coefficients towards zero. We leave the analysis of such open problems for future research.

Acknowledgments

We thank the Editor, an Associate Editor, and two Reviewers for their fruitful comments and helpful

suggestions on an earlier version of this paper. We thank, without implicating, I. Papastathopoulos, R. Huser, and V. Inácio de Carvalho for a variety of helpful comments and feedback, and L. Mhalla for insightful discussions and for sharing the R code for fitting the methods of Mhalla et al. (2019a). MdC was partially supported by FCT (Fundação para a Ciência e a Tecnologia, Portugal) through the projects PTDC/MAT-STA/28649/2017 and UID/MAT/00006/2020.

References

- Albuquerque, R. and Vega, C. (2009) Economic news and international stock market co-movement. *Review of Finance*, **13**, 401–465.
- Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004) *Statistics of Extremes: Theory and Applications*. Wiley. Hoboken, NJ: Wiley.
- Berman, S. M. (1961) Convergence to bivariate limiting extreme value distributions. *Ann. Inst. Statist. Math*, **13**, 211–223.
- de Boor, C. (2001) *A Practical Guide to Splines*. New York: Springer.
- Brezger, A. and Lang, S. (2006) Generalized structured additive regression based on Bayesian P-splines. *Computational Statistics & Data Analysis*, **50**, 967–991.
- Brooks, R. and Del Negro, M. (2004) The rise in comovement across national stock markets: Market integration or it bubble? *Journal of Empirical Finance*, **11**, 659–680.
- Bruno, F., Greco, F. and Ventrucci, M. (2016) Non-parametric regression on compositional covariates using Bayesian P-splines. *Statistical Methods & Applications*, **25**, 75–88.
- Castro, D., de Carvalho, M. and Wadsworth, J. L. (2018) Time-varying extreme value dependence with application to leading European stock markets. *The Annals of Applied Statistics*, **12**, 283–309.
- Chirac, J. (2006) European Council Statements. Brussels, 24 March 2006.
- Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values*. London: Springer London.
- Coles, S., Heffernan, J. and Tawn, J. (1999) Dependence measures for extreme value analyses. *Extremes*, **2**, 339–365. URL: <https://doi.org/10.1023/A:1009963131610>.
- de Carvalho, M. (2016) Statistics of extremes: Challenges and opportunities. In *Extreme Events in Finance: A Handbook of Extreme Value Theory and Its Applications* (ed. F. Longin). Hoboken: Wiley.
- Dobson, A. J. (2008) *An Introduction to Generalized Linear Models*. Chapman & Hall/CRC. Boca Raton, FL: Chapman & Hall/CRC.
- Ehrmann, M. and Jansen, D.-J. (2020) Stock return comovement when investors are distracted: More, and more homogeneous. *European Central Bank Working Paper*, No 2412.
- Eilers, P. H. and Marx, B. D. (2021) *Practical Smoothing: The Joys of P-splines*. Cambridge: Cambridge University Press.
- Embrechts, P., Hofert, M. and Wang, R. (2016) Bernoulli and tail-dependence compatibility. *The Annals of Applied Probability*, **26**, 1636–1658.
- Embrechts, P., McNeil, A. and Straumann, D. (2002) *Risk Management: Value at Risk and Beyond*, chap. Correlation and Dependence in Risk Management: Properties and Pitfalls, 176–223. Cambridge: Cambridge University Press.
- Ergen, I. (2014) Tail dependence and diversification benefits in emerging market stocks: An extreme value theory approach. *Applied Economics*, **46**, 2215–2227.
- Escobar-Bach, M., Goegebeur, Y. and Guillou, A. (2018) Local robust estimation of the Pickands dependence function. *The Annals of Statistics*, **46**, 2806–2843.
- Fahrmeir, L., Kneib, T. et al. (2011) *Bayesian Smoothing and Regression for Longitudinal, Spatial and Event History Data*. Oxford: Oxford University Press.
- Ferrez, J., Davison, A. and Rebetez, M. (2011) Extreme temperature analysis under forest cover compared to an open field. *Agricultural and Forest Meteorology*, **151**, 992–1001.
- Forbes, K. J. and Rigobon, R. (2002) No contagion, only interdependence: Measuring stock market comovements. *The Journal of Finance*, **57**, 2223–2261.
- Gentle, J. E. (2020) *Statistical Analysis of Financial Data: With Examples in R*. Boca Raton, FL: Chapman & Hall/CRC.

- Gong, Y. and Huser, R. (2022) Asymmetric tail dependence modeling, with application to cryptocurrency market data. *The Annals of Applied Statistics*, **16**, 1822–1847.
- Guette-Khiter, C. (2016) Non-resident holdings of french CAC 40 companies at end-2015. *Quarterly Selection of Articles—Bulletin de la Banque de France*, 35–46.
- Jach, A. (2017) International stock market comovement in time and scale outlined with a thick pen. *Journal of Empirical Finance*, **43**, 115–129.
- Koop, G. (2006) *Analysis of Financial Data*. Chichester: John Wiley & Sons.
- Lang, S. and Brezger, A. (2004) Bayesian P-splines. *Journal of Computational and Graphical Statistics*, **13**, 183–212.
- Ledford, A. W. and Tawn, J. A. (1996) Statistics for near independence in multivariate extreme values. *Biometrika*, **83**, 169–187.
- (1997) Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **59**, 475–499.
- Longin, F. and Solnik, B. (2001) Extreme correlation of international equity markets. *The Journal of Finance*, **56**, 649–676.
- Marcon, G., Padoan, S. A. and Antoniano-Villalobos, I. (2016) Bayesian inference for the extremal dependence. *Electronic Journal of Statistics*, **10**, 3310–3337.
- McCullagh, P. and Nelder, J. A. (1989) *Generalized Linear Models*. Boca Raton, FL: Chapman & Hall/CRC.
- Mhalla, L., Chavez-Demoulin, V. and Naveau, P. (2017) Non-linear models for extremal dependence. *Journal of Multivariate Analysis*, **159**, 49–66.
- Mhalla, L., de Carvalho, M. and Chavez-Demoulin, V. (2019a) Regression type models for extremal dependence. *Scandinavian Journal of Statistics*, **46**, 1141–1167.
- Mhalla, L., Opitz, T. and Chavez-Demoulin, V. (2019b) Exceedance-based nonlinear regression of tail dependence. *Extremes*, **22**, 523–552.
- Morana, C. and Beltratti, A. (2008) Comovements in international stock markets. *Journal of International Financial Markets, Institutions and Money*, **18**, 31–45.
- Patton, A. J. (2006) Modelling asymmetric exchange rate dependence. *International Economic Review*, **47**, 527–556.
- Poon, S.-H., Rockinger, M. and Tawn, J. (2003) Modelling extreme-value dependence in international stock markets. *Statistica Sinica*, 929–953.
- (2004) Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *Review of Financial Studies*, **17**, 581–610.
- Quintos, C., Fan, Z. and Phillips, P. C. (2001) Structural change tests in tail behaviour and the Asian crisis. *The Review of Economic Studies*, **68**, 633–663.
- Rondon, L. M. and Bolfarine, H. (2016) Bayesian analysis of generalized elliptical semi-parametric models. *Journal of Applied Statistics*, **43**, 1508–1524.
- Rua, A. and Nunes, L. C. (2009) International comovement of stock market returns: A wavelet analysis. *Journal of Empirical Finance*, **16**, 632–639.
- Ruppert, D. (2002) Selecting the number of knots for penalized splines. *Journal of computational and graphical statistics*, **11**, 735–757.
- Sriram, K., Shi, P. and Ghosh, P. (2016) A Bayesian quantile regression model for insurance company costs data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **179**, 177–202.
- Statista (2021) Distribution of countries with largest stock markets worldwide as of January 2021, by share of total world equity market value. *Statista Research Department*, Source: Credit Suisse; MSCI; FTSE; S&P Global.
- Wang, H. and Tsai, C.-L. (2009) Tail index regression. *Journal of the American Statistical Association*, **104**, 1233–1240.
- Wood, S. (2017) *Generalized Additive Models: An Introduction with R*. Boca Raton: Chapman & Hall/CRC.
- Yuan, M. and Lin, Y. (2006) Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **68**, 49–67.