Supporting information for: 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

1. Posterior sampling from χ_t and $\bar{\chi}_t$

In this section we provide details on posterior sampling for the proposed models. Since the pseudo-observations I_t and E_t are asymptotically (as $u \to \infty$) Bernoulli and Exponential-distributed, inference can be framed into a framework of GLMs (Generalized Linear Models).

1.1. A Metropolis–Hastings algorithm for Bayesian P-splines with IWLS proposals

Both χ_t and $\bar{\chi}_t$ can be fitted using a Metropolis–Hastings algorithm for Bayesian P-splines with IWLS (Iteratively Weighted Least Squares) proposals. Similar algorithms can be found in Fahrmeir et al. (2011, Chapter 2). We first focus on χ_t. Let $\{(T_i, I_i)\}_{i=1}^{k_I}$ be the pseudo-observations from which the estimate of χ_t is to be produced, where the times of the exceedances are $\{T_1, \dots, T_{k_I}\} = \{t : Y_t > u\}$. Define the weight matrix $\mathbf{W}$ as the $k_I \times k_I$ diagonal matrix with elements

$$w_{i,i} = \frac{[F'\{g(T_i)\}]^2}{F\{g(T_i)\}[1 - F\{g(T_i)\}]}, \tag{1}$$

where $F'(x) = \mathrm{d}F/\mathrm{d}x$, and define the working responses as the k_I-vector, $\mathbf{z} = (z_1, \dots, z_{k_I})^{\mathrm{T}}$, with elements

$$z_i = g(T_i) + \frac{I_i - F\{g(T_i)\}}{F'\{g(T_i)\}}. \tag{2}$$

For example, if $\chi_t = F\{g(t)\}$ with $F(x) = \exp(x)/\{1 + \exp(x)\}$, then

$$w_{i,i} = \frac{\exp\{g(T_i)\}}{[1 + \exp\{g(T_i)\}]^2}, \qquad z_i = g(T_i) + \left[I_i - \frac{1}{1 + \exp\{-g(T_i)\}}\right] \frac{\exp\{g(T_i)\}}{[1 + \exp\{g(T_i)\}]^2}.$$

Below, the superscript 'c' denotes current, or based on current, 'p' stands for proposal, and $\Phi(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the distribution function of the multivariate Normal distribution, with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. As in the paper, below $\mathbf{K}$ denotes the penalty matrix.

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

Algorithm 1 Metropolis–Hastings with IWLS Proposals (for χ_t and $\bar{\chi}_t$)

1) Update $\boldsymbol{\beta}$: Draw a proposal for $\boldsymbol{\beta}^p$ from a multivariate Gaussian

$$\mathrm{N}(\boldsymbol{\mu}^c, \boldsymbol{\Sigma}^c),$$

where $\boldsymbol{\mu}^c \equiv \boldsymbol{\Sigma}^c \mathbf{X}^{\mathrm{T}} \mathbf{W}^c \mathbf{z}^c$ and $\boldsymbol{\Sigma}^c \equiv \{\mathbf{X}^{\mathrm{T}} \mathbf{W}^c \mathbf{X} + (\tau^{-2})^c \mathbf{K}\}^{-1}$, and accept it with probability,

$$\alpha(\boldsymbol{\beta}^c, \boldsymbol{\beta}^p) = \frac{\mathsf{p}(\mathbf{y} \mid \boldsymbol{\beta}^p)\pi\{\boldsymbol{\beta}^p \mid (\tau^2)^c\}\Phi(\boldsymbol{\beta}^c; \boldsymbol{\mu}^p, \boldsymbol{\Sigma}^p)}{\mathsf{p}(\mathbf{y} \mid \boldsymbol{\beta}^c)\pi\{\boldsymbol{\beta}^c \mid (\tau^2)^c\}\Phi(\boldsymbol{\beta}^p; \boldsymbol{\mu}^c, \boldsymbol{\Sigma}^c)}.$$

2) Update τ^2: Draw a new $(\tau^2)^c$ from the Inverse Gamma distribution,

$$\mathrm{IG}(a + \mathrm{rank}(\mathbf{K})/2, b + 1/2\, \boldsymbol{\beta}^{\mathrm{T}} \mathbf{K} \boldsymbol{\beta}).$$

Algorithm 1 can be used for fitting χ_t along with any inverse link function F, by setting $\mathbf{y} = (I_1, \dots, I_{k_I})^{\mathrm{T}}$ and

$$\mathbf{X} = \begin{pmatrix} B_1^d(T_1) & \cdots & B_K^d(T_1) \\ \vdots & \vdots & \vdots \\ B_1^d(T_{k_I}) & \cdots & B_K^d(T_{k_I}) \end{pmatrix}. \tag{3}$$

The same algorithm can also be used for fitting $\bar{\chi}_t$ by adjusting the definitions of weight matrix and working observations from Equations (1) and (2), and by setting $\mathbf{y} = (E_1, \dots, E_{k_E})^{\mathrm{T}}$ and

$$\mathbf{X} = \begin{pmatrix} B_1^d(T_1') & \cdots & B_K^d(T_1') \\ \vdots & \vdots & \vdots \\ B_1^d(T_{k_E}') & \cdots & B_K^d(T_{k_E}') \end{pmatrix},$$

with $\{T_1', \dots, T_{k_E}'\} = \{t : Z_t > u\}$. For $\bar{\chi}_t$ the weight matrix is a $k_E \times k_E$ diagonal matrix and the working response is a k_E-vector, $\mathbf{z} = (z_1, \dots, z_{k_E})^{\mathrm{T}}$, whose respective elements are:

$$w_{i,i} = \left(\frac{H'\{l(T_i')\}}{H\{l(T_i')\}}\right)^2, \qquad z_i = l(T_i') + \frac{E_i - H\{l(T_i')\}}{H'\{l(T_i')\}}.$$

For example, when $\bar{\chi}_t = 2H\{l(T_i')\} - 1$ with $H(x) = \Phi(x)$, then

$$w_{i,i} = \left(\frac{\phi\{l(T_i')\}}{\Phi\{l(T_i')\}}\right)^2, \qquad z_i = l(T_i') + \frac{E_i - \Phi\{l(T_i')\}}{\phi\{l(T_i')\}},$$

where $\Phi(x)$ and $\phi(x)$ are respectively the distribution function and the density of the standard Normal distribution. Next, we note that χ_t can be fitted using a Gibbs sampler, for a specific instance of the link function.

Algorithm 2 Gibbs Sampler (for χ_t with link function $F(x) = \Phi^{-1}(x)$)

1) Update latent variables: Draw $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)^{\mathrm{T}}$ from

$$\lambda_i \mid \text{else} \sim \begin{cases} \mathrm{TN}_{[0,\infty)}(F\{g(T_i)\}, 1), & I_i = 1, \\ \mathrm{TN}_{[-\infty,0)}(F\{g(T_i)\}, 1), & I_i = 0. \end{cases}$$

2) Update $\boldsymbol{\beta}$: Draw $\boldsymbol{\beta}$ from $\mathrm{N}(\boldsymbol{\Sigma}\mathbf{X}^{\mathrm{T}}\boldsymbol{\lambda}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = (\mathbf{X}^{\mathrm{T}}\mathbf{X} + \tau^{-2}\mathbf{K})^{-1}$.
3) Update τ^2: Draw τ^2 from $\mathrm{IG}(a + \mathrm{rank}(\mathbf{K})/2, b + 1/2\, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{K}\boldsymbol{\beta})$.

Table 1. MIAE (Mean Integrated Absolute Error) from the Monte Carlo simulation study for posterior mean time-varying extremal dependence measures χ_t and $\bar{\chi}_t$.

sample size	χ_t				$\bar{\chi}_t$			
	A	B	C	D	A	B	C	D
10 000	0.0485	0.0355	0.0135	0.0332	0.0218	0.0169	0.2308	0.2015
20 000	0.0370	0.0281	0.0109	0.0260	0.0138	0.0122	0.1380	0.1440
40 000	0.0280	0.0215	0.0088	0.0194	0.0108	0.0079	0.1020	0.1130

1.2. A Gibbs sampler for χ_t

When the inverse link function for χ_t is $F(x) = \Phi(x)$, then fitting χ_t boils down to fitting a standard logistic regression model to the pseudo-observations $\{(T_i, I_i)\}_{i=1}^{k_I}$. Hence, in that case inference for χ_t can be conducted using a well-known latent specification due to Albert and Chib (1993), which yields a Gibbs sampler. Specifically, consider the latent Gaussian variable,

$$\lambda_i = F\{g(T_i)\} + \varepsilon_i, \qquad \varepsilon_i \sim N(0, 1),$$

so that $I_i = 1$ if and only if $\lambda_i \geq 0$; let $\mathrm{TN}_{[a,b]}(\mu, \sigma^2)$ be the Normal distribution with mean μ and variance σ^2, truncated to $[a, b]$. The resulting Gibbs sampler is in Algorithm 2, with $\mathbf{X}$ as in (3).

2. Supplementary materials for numerical studies

This section supplements the numerical studies from Section 3 of the paper.

2.1. Time-invariant margins

Fig. 1 shows the trajectories of the posterior means for 100 randomly selected samples along with the Monte Carlo mean from the simulation study from Section 3.2 of the paper. As it can be seen from Fig. 1 the proposed methods satisfactorily recover the true χ_t and $\overline{\chi}_t$.

We have also examined the frequentist behaviour of the methods from a numerical stance, by computing the MIAE (Mean Integrated Absolute Error) over different sample sizes As expected, for both measures of time-changing extremal dependence, χ_t and $\bar{\chi}_t$, the MIAE reduces significantly as the sample size increases—as can be seen from Table 1.

In addition, we have also compared the proposed methods with the exceedance-based regression methods of Mhalla et al. (2019). Fig. 2 depicts the trajectories of χ_t and $\bar{\chi}_t$ estimated by the exceedance-based regression methods of (Mhalla et al., 2019), obtained from 100 randomly selected samples of the Monte Carlo simulation. While the estimates based on the methods of Mhalla et al. are reasonably on target, they tend to be more wiggly than those of the proposed methods in terms of χ_t, and $\bar{\chi}_t$ tends to be more biased under asymptotic dependence. For Scenarios D the fits of $\bar{\chi}_t$ are slightly better for the method of Mhalla et al..

Finally, following a recommendation by a reviewer, we also conducted additional numerical experiments with an inverted extreme value distribution (Ledford and Tawn, 1997). Namely, we consider an a time-varying inverted logistic distribution with Laplace margins, having joint survivor function

$$P(X_t > x, Y_t > x) = \exp\left[-V_t\left(\frac{-1}{\log[1/\{2\exp(-x)\}]}, \frac{-1}{\log[1/\{2\exp(-y)\}]}\right)\right], \quad x, y > 0,$$

and where $V_t(a, b) = (a^{-1/\theta_t} + b^{-1/\theta_t})^{\theta_t}$ is the corresponding exponent function, with the same θ_t in Section 3. As it can be seen from Fig. 3 this scenario generates 'strong' asymptotic independence (i.e., large values of η_t) and hence not surprisingly the sub-asymptotic $\chi_t(u) = P(X_t > u \mid Y_t > u)$ is far from its limiting value zero. Still, similarly to Scenario D, the fitted χ_t satisfactorily recovers its sub-asymptotic version.

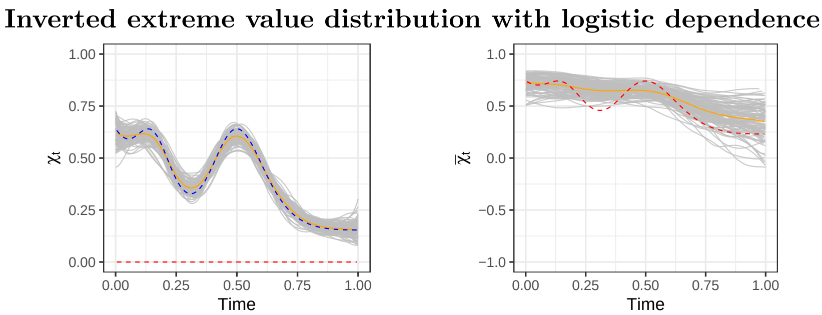

Fig. 3. Additional Monte Carlo simulation for the proposed method for a scenario based on an inverted extreme value distribution (time-invariant margins): Monte Carlo mean (solid orange line), true (red dashed line), and the true sub-asymptotic $\chi_t(u)$ (blue dashed line). 100 randomly selected trajectories of posterior means are depicted in light grey.

2.2. *Time-varying margins*

We now report supplementary Monte Carlo experiments considering the same dependence structure as in Scenarios A–D in the paper but with time-varying margins. We refer to these novel simulation setups as Scenarios I–IV and margins change over time as follows, $X_t \sim \text{GEV}(\mu_t^X, \sigma_t^X, 1)$ and $Y_t \sim \text{GEV}(\mu_t^Y, \sigma_t^Y, 1)$, with

$$(\mu_t^X, \sigma_t^X) = (2\sin(3t) + 10, 2 + 3t/2), \quad (\mu_t^X, \sigma_t^X) = (5\sin(10t - 3) + 5, 2\cos(5t)/2 + 1.5).$$

As it can be seen from Fig. 4, the performance of the proposed methods is still quite satisfactory even in this case. Some comments on implementations are as in order. We transform the simulated (X_t, Y_t) to unit Fréchet margins $(\mathcal{X}_t, \mathcal{Y}_t)$ using the transformation,

$$(\mathcal{X}_t, \mathcal{Y}_t) = (-1/\log\{F_{X_t}(X_t)\}, -1/\log\{F_{Y_t}(Y_t)\}),$$

where F_{X_t} and F_{Y_t} are the respective marginal time-varying distribution functions for X_t and Y_t, which are fitted using the time-varying distribution function estimator of Harvey and Oryshchenko (2012).

3. Supplementary materials for real data application

3.1. *Summary statistics and additional pairs of stocks*

Table 2 presents summary statistics for the six stock index returns under study. All returns show evidence of negative skewness and a kurtosis significantly greater than three; such empirical attributes are well-known and are often called stylized facts (e.g., Gentle, 2020, Section 1.6).

Fig. 5 presents the patterns of extremal dependence of stock indices obtained by fitting χ_t and $\bar{\chi}_t$, for pairs which were not covered in the main article.

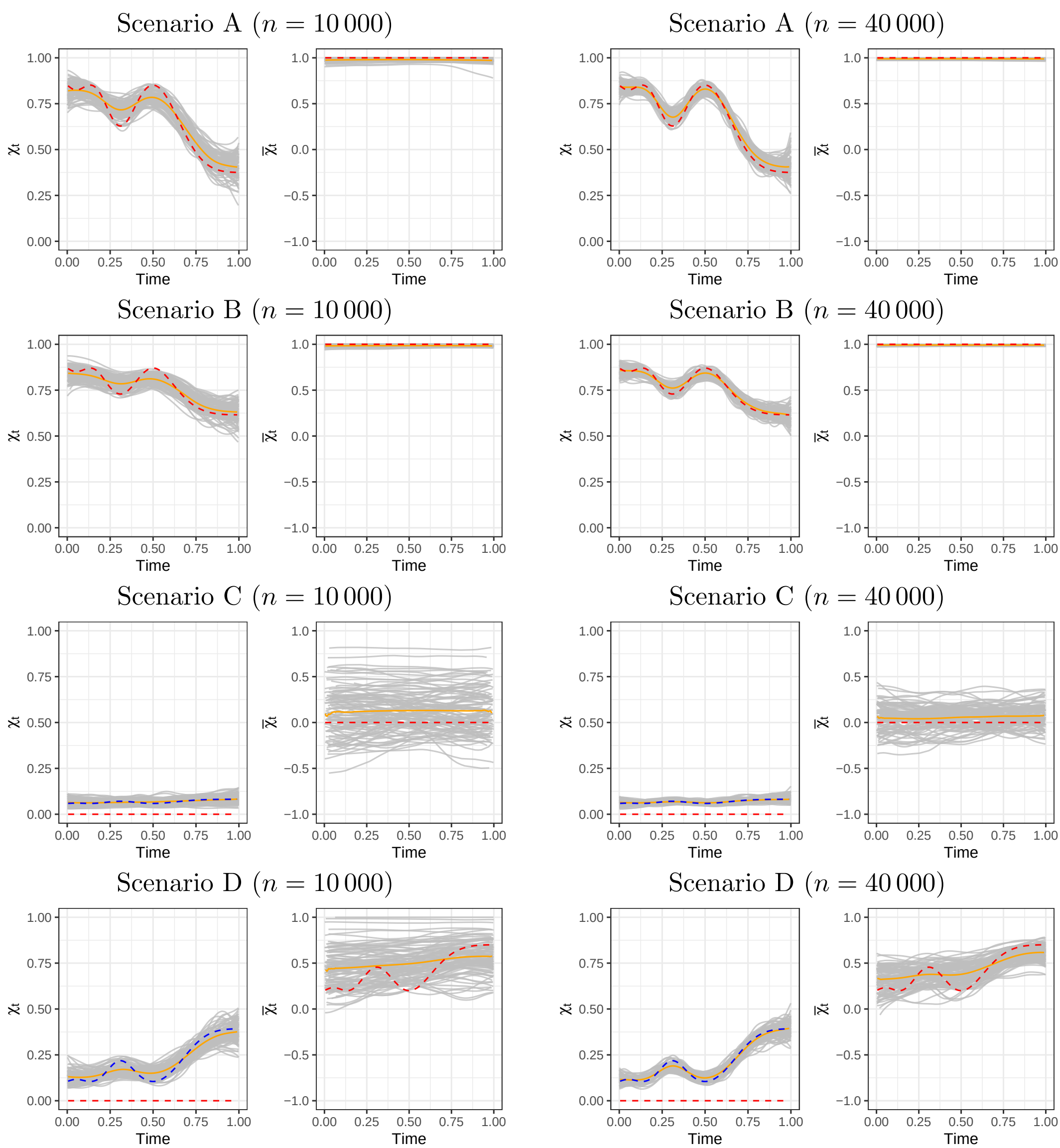

Fig. 1. Monte Carlo simulation for the proposed method (time-invariant margins): Monte Carlo mean (solid orange line), true (red dashed line), and the true sub-asymptotic $\chi_t(u)$ (blue dashed line). 100 randomly selected trajectories of posterior means are depicted in light grey.

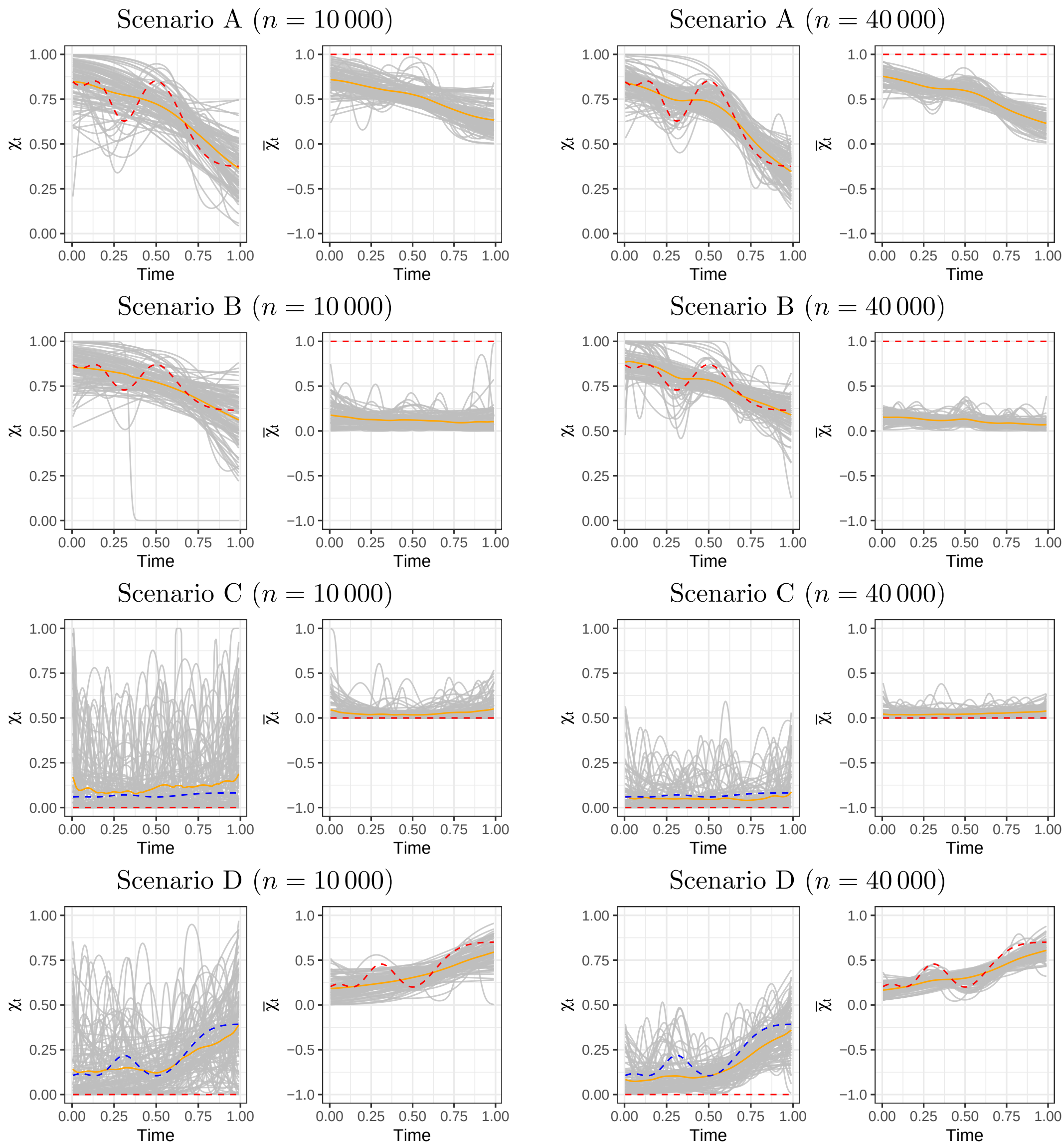

Fig. 2. Monte Carlo simulation for the method of Mhalla et al. (time-invariant margins): Monte Carlo mean (solid orange line), true (red dashed line), and the true sub-asymptotic $\chi_t(u)$ (blue dashed line). 100 randomly selected trajectories of the fitted lines are depicted in light grey.

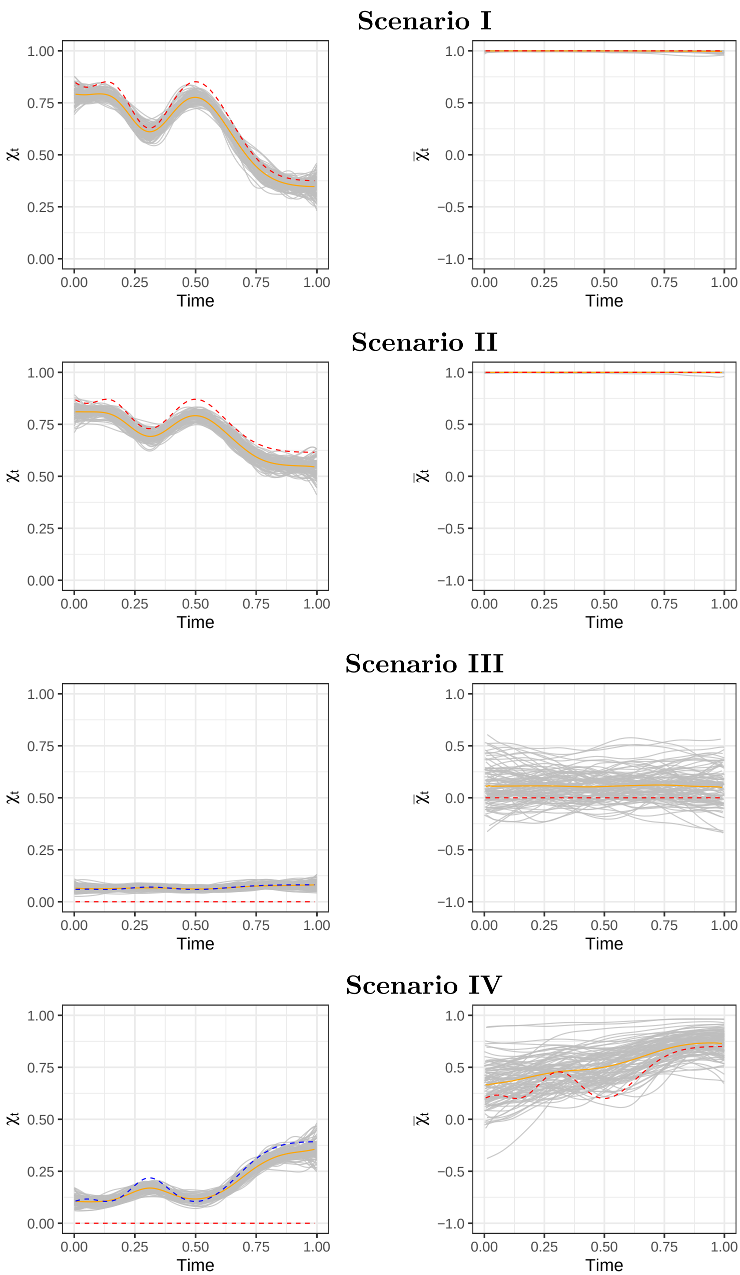

Fig. 4. Monte Carlo simulation for the proposed method (time-varying margins): Monte Carlo mean (solid orange line) from simulation study in Section 3 plotted against the values from true dual measures (red dashed line) and the true sub-asymptotic $\chi_t(u)$ (blue dashed line). 100 randomly selected trajectories of posterior means are depicted in light grey.

Table 2. Summary statistics of stock index returns

	UK	**FRA**	**GER**	**CHN**	**JPN**	**US**
Mean	0.0002	0.0001	0.0002	0.0003	0.0002	0.0003
S.D.	0.0110	0.0138	0.0141	0.0108	0.0130	0.0158
Skewness	-0.5795	-0.2075	-0.3310	-1.0220	-0.4104	-0.6685
Kurtosis	10.5966	5.7852	6.9206	25.6064	9.4108	16.1292

3.2. Sensitivity analysis

Figures 6 and 7 present sensitivity analyses of the results presented in Section 5 of the paper, with $m + 1 = 30$ knots and degree $d = 4$; the key empirical findings are tantamount to the ones reported in the paper.

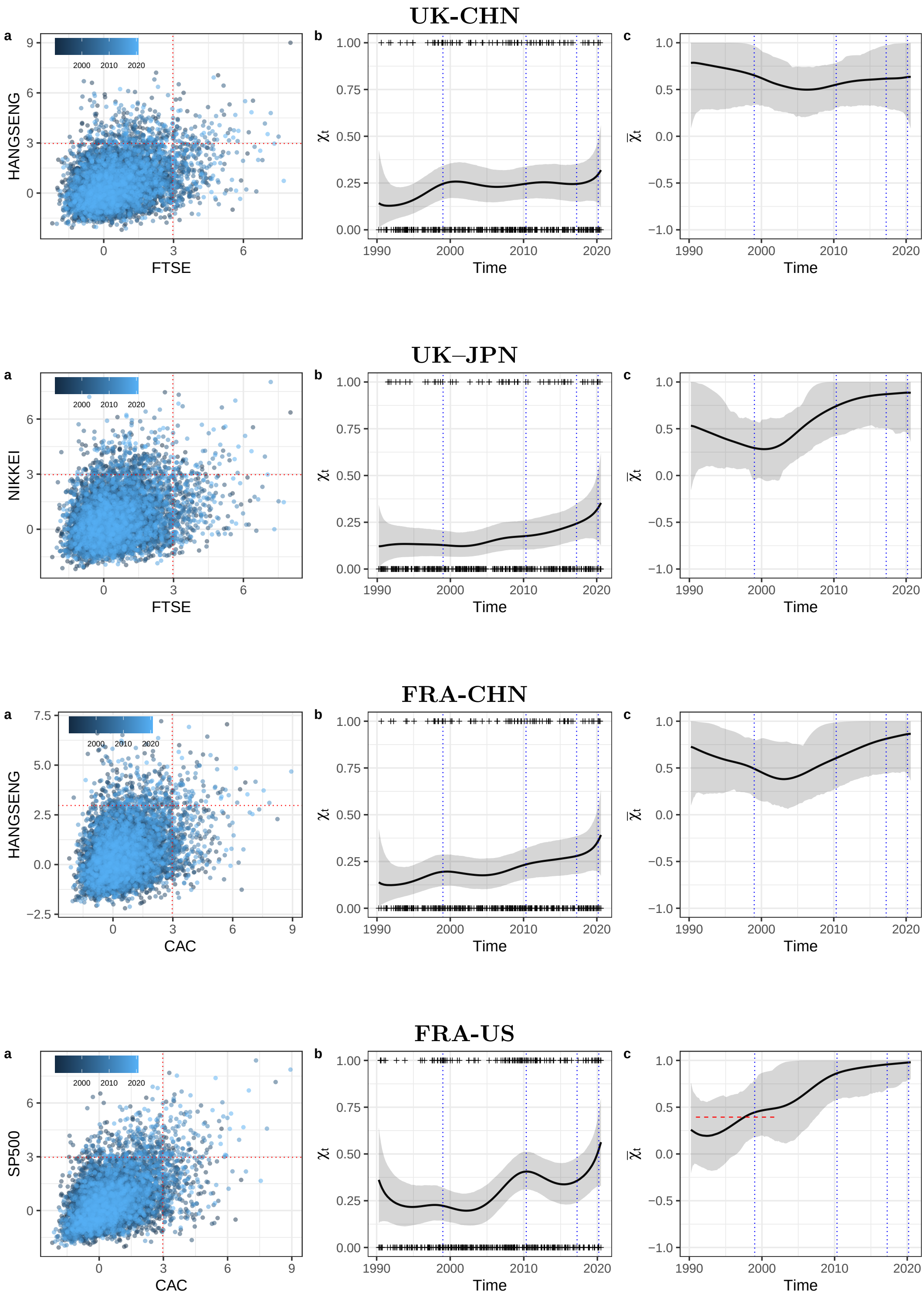

Fig. 5. Supplementary between-regions and with US analyses. Left: Scatterplot of log transformed data. Middle and Right: Posterior mean time-varying χ_t and $\bar{\chi}_t$ (black solid) along with credible bands (grey area); the rug in the middle panel corresponds to the points $\{(t, \mathbb{1}_{\{X_t>u\}}) : Y_t > u\}$.

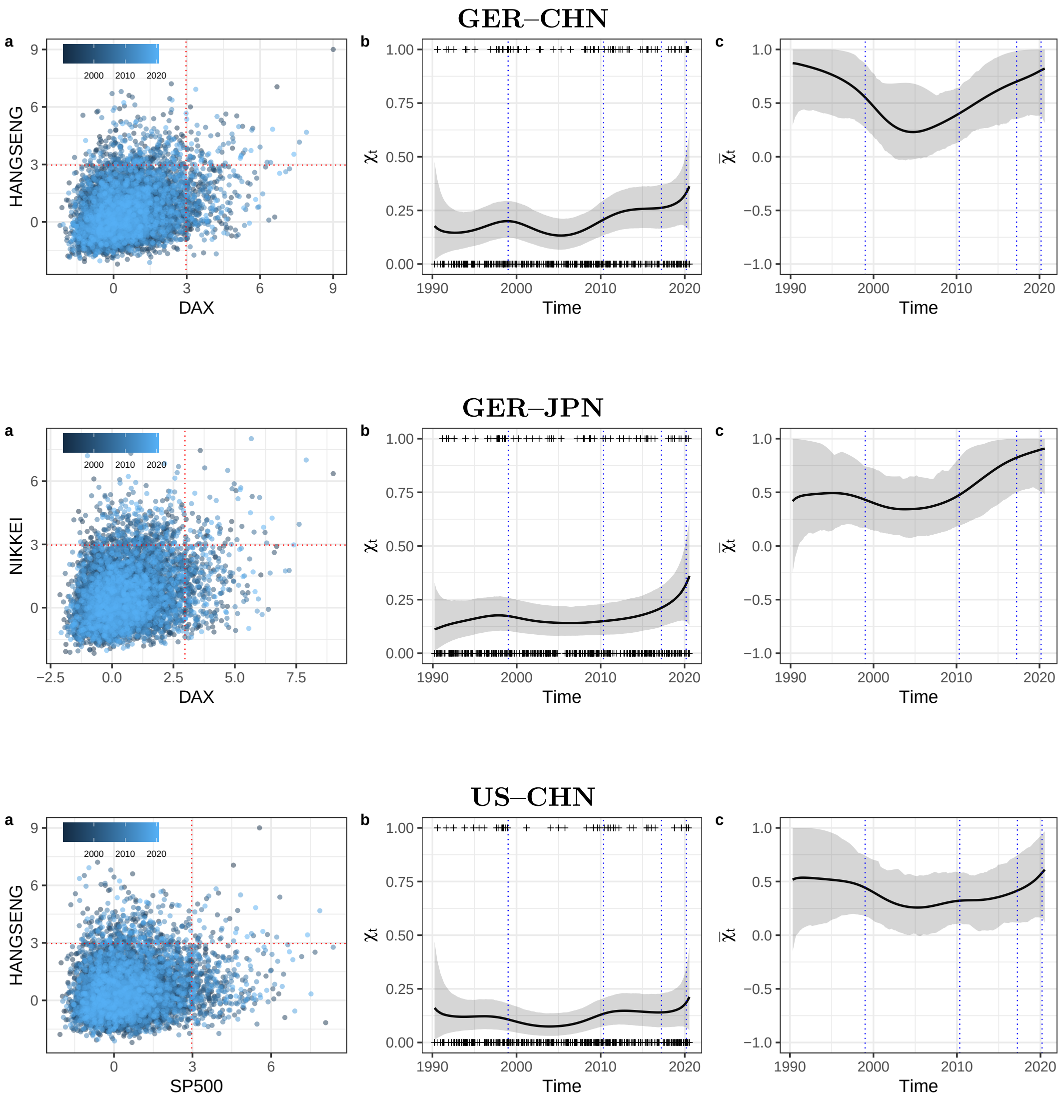

Fig. 5. Supplementary between-regions and with US analyses (cont). Left: Scatterplot of log transformed data. Middle and Right: Posterior mean time-varying χ_t and $\bar{\chi}_t$ (black solid) along with credible bands (grey area); the rug in the middle panel corresponds to the points $\{(t, \mathbb{1}_{\{X_t>u\}}) : Y_t > u\}$.

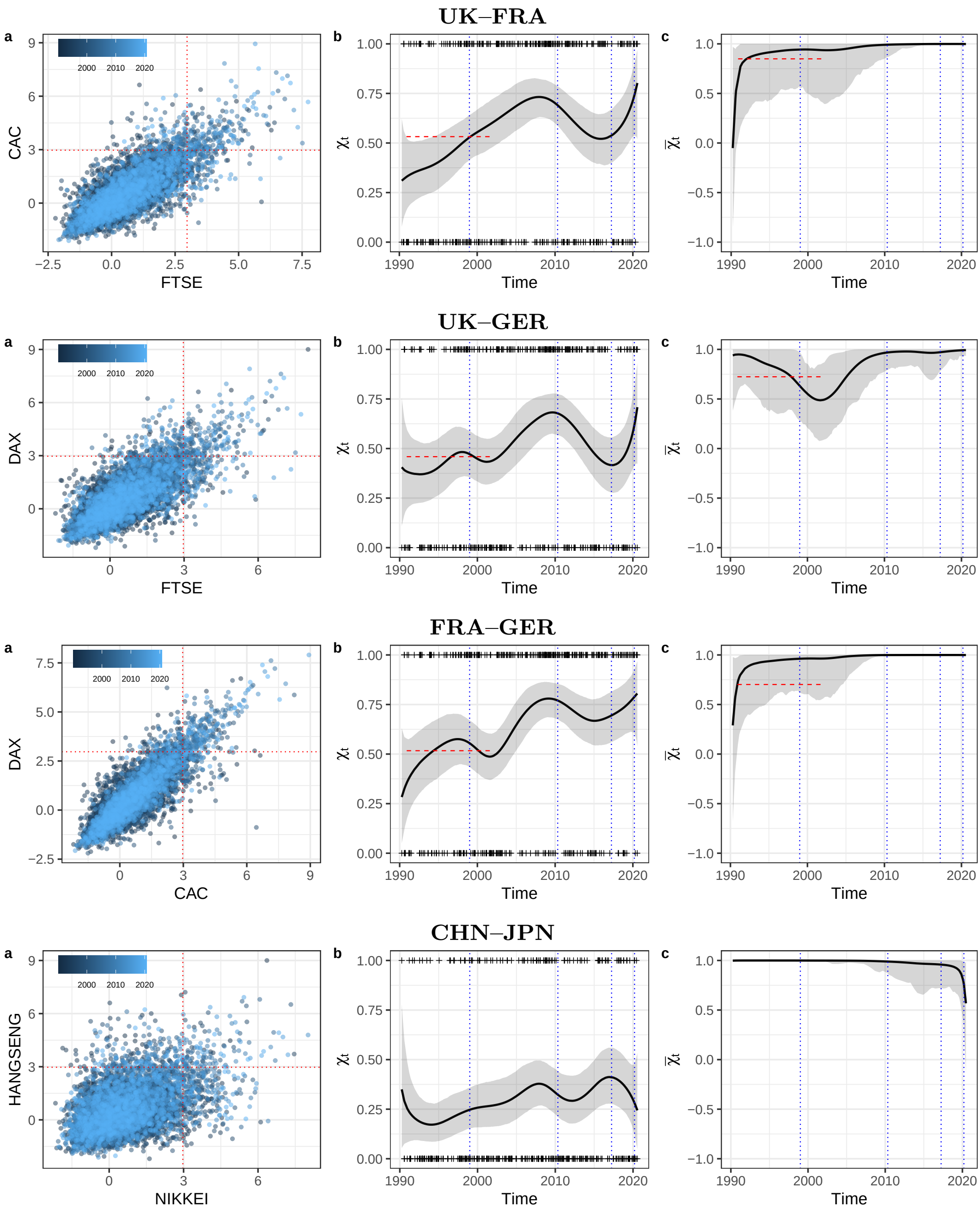

Fig. 6. Sensitivity analysis—within-region. Left: Scatterplot of log transformed data. Middle and Right: Posterior mean time-varying χ_t and $\bar{\chi}_t$ (solid) along with credible bands; the rug in the middle panel corresponds to the points $\{(t, \mathbb{1}_{\{X_t>u\}}) : Y_t > u\}$ whereas the 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).

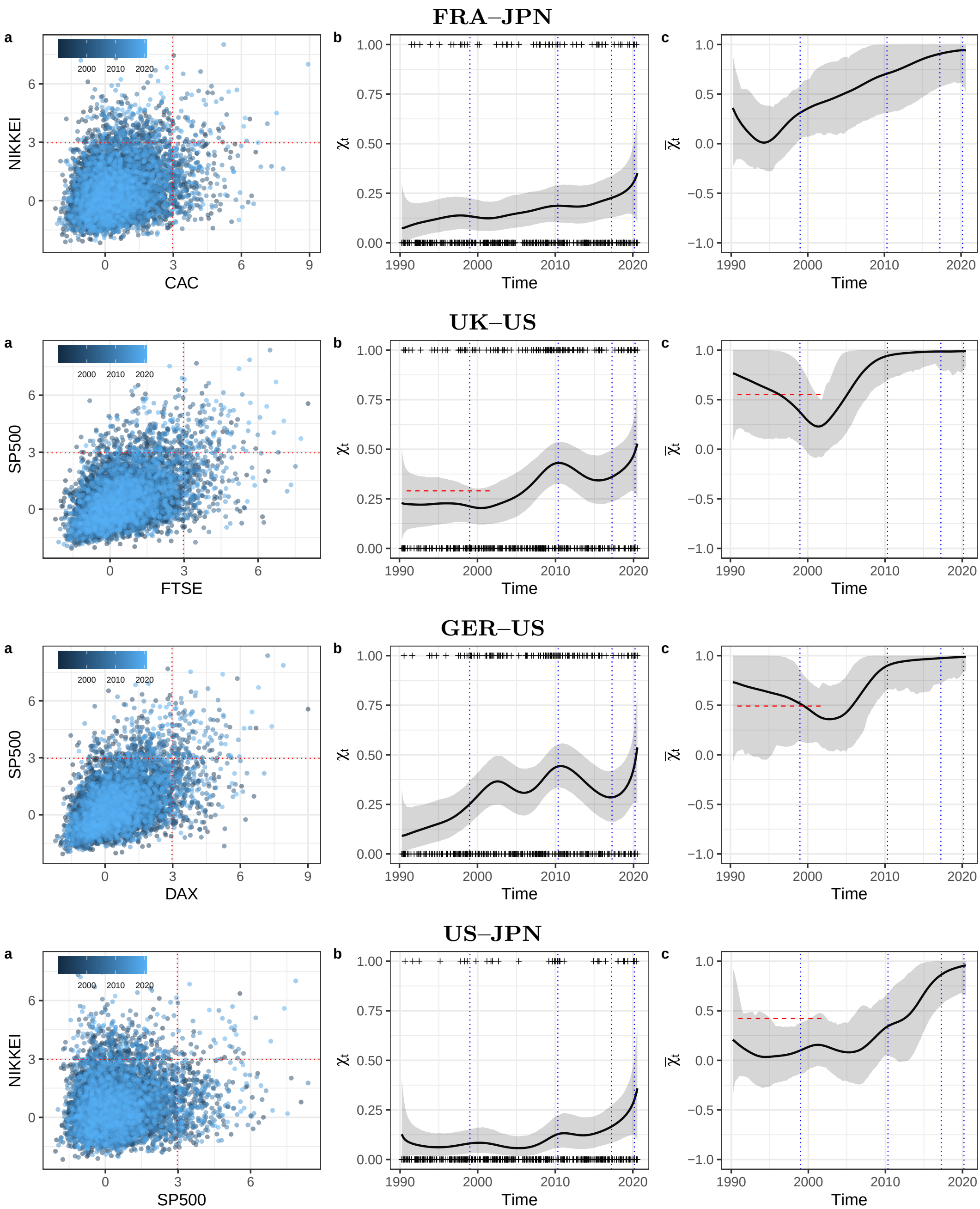

Fig. 7. Sensitivity analysis—between-regions and with US. Left: Scatterplot of log transformed data. Middle and Right: Posterior mean time-varying χ_t and $\bar{\chi}_t$ (black solid) along with credible bands (grey area); the rug in the middle panel corresponds to the points $\{(t, \mathbb{1}_{\{X_t>u\}}) : Y_t > u\}$ whereas the dashed line corresponds to the available values from the subperiod analysis of Poon et al. (2003).

References

Albert, J. H. and Chib, S. (1993) Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, **88**, 669.

Fahrmeir, L., Kneib, T. et al. (2011) *Bayesian Smoothing and Regression for Longitudinal, Spatial and Event History Data*. Oxford: Oxford University Press.

Gentle, J. E. (2020) *Statistical Analysis of Financial Data: With Examples in R*. Boca Raton, FL: Chapman & Hall/CRC.

Harvey, A. and Oryshchenko, V. (2012) Kernel density estimation for time series data. *International journal of forecasting*, **28**, 3–14.

Ledford, A. W. and Tawn, J. A. (1997) Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **59**, 475–499.

Mhalla, L., Opitz, T. and Chavez-Demoulin, V. (2019) Exceedance-based nonlinear regression of tail dependence. *Extremes*, **22**, 523–552.

Poon, S.-H., Rockinger, M. and Tawn, J. (2003) Modelling extreme-value dependence in international stock markets. *Statistica Sinica*, 929–953.