Supporting information for:
Semiparametric Bayesian modeling of nonstationary joint extremes: How do big tech's extreme losses behave?

Miguel de Carvalho[1,2] and Karla Vianey Palacios Ramirez[3]

[1] *School of Mathematics, University of Edinburgh, Edinburgh, UK*
[2] *Department of Mathematics–CIDMA, Universidade de Aveiro, Aveiro, Portugal*
[3] *School of Mathematics, Statistics and Physics, Newcastle University, Newcastle, UK*

1. Supplementary simulation studies

1.1. EDI kernel density estimator

The Monte Carlo experiments to be reported next illustrate: i) the boundary-bias issue of the EDI kernel density estimator; ii) superior performance of the Polya tree-based EDI over the kernel approach. The EDI kernel density estimator is

$$\hat{f}(t) = \frac{1}{k}\sum_{j=1}^{k} K_h\left(t - \tau_j\right), \tag{1}$$

where $h > 0$ is the bandwidth and $K_h(\cdot) = K(\cdot/h)/h$, with K denoting a kernel. Fig. 1 emphasizes the well-known boundary-bias issue of kernel density estimators. For the experiments in Fig. 1, we use $h = 0.1$ along with an Epanechnikov kernel which is known to be optimal under mild conditions (Wand and Jones, 1995, Section 2.7); a similar poor fit at the boundary is visible using other bandwidth selection methods such as, for example, Silverman's rule of thumb (Silverman, 1986).

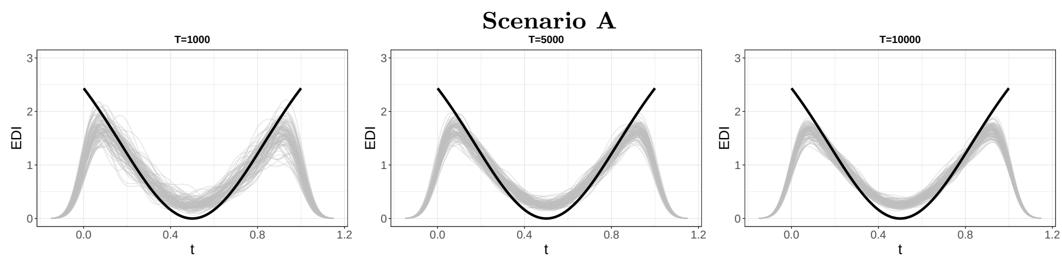

Fig. 1. Boundary-bias issue of the EDI kernel density estimator.

To compare the performance of the EDI Polya tree-based estimator versus the kernel approach in (1), we compute the MISE (Mean Integrated Squared Error) for Scenarios A–C under the different sample sizes. Table 1 reports the results and it suggests that the mixture of finite Polya trees method provides a better fit in general.

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

Table 1. Monte Carlo mean MISE for Scenarios A–C for EDI Polya tree-based estimator versus kernel density estimator.

Scenario A	$T = 500$	$T = 1\,000$	$T = 5\,000$	$T = 10\,000$
Polya	0.0207	0.0182	0.0120	0.0103
Kernel	0.2575	0.1390	0.0928	0.0853
Scenario B	$T = 500$	$T = 1\,000$	$T = 5\,000$	$T = 10\,000$
Polya	0.0152	0.0106	0.0070	0.0057
Kernel	0.0261	0.0229	0.0078	0.0067
Scenario C	$T = 500$	$T = 1\,000$	$T = 5\,000$	$T = 10\,000$
Polya	0.0265	0.0126	0.0089	0.0080
Kernel	0.0272	0.0186	0.0090	0.0086

Table 2. Monte Carlo posterior median coefficient of tail dependence and coverage probabilities I_x at levels $x = 0.90, 0.95$. The true coefficients of tail dependence for Scenarios D and E are γ = 0.3 and $\gamma = 0.5$, respectively.

Scenario	**Sample size** (T)	**Monte Carlo posterior median**	$I_{.90}$	$I_{.95}$
D	500	0.3039	0.810	0.910
E	500	0.5048	0.802	0.896
D	1 000	0.3037	0.860	0.970
E	1 000	0.5027	0.864	0.908
D	5 000	0.3001	0.897	0.983
E	5 000	0.5000	0.900	0.959
D	10 000	0.3000	0.924	0.999
E	10 000	0.5000	0.910	0.969

1.2. Asymptotic independence

This section considers two additional simulation scenarios with $0 < \gamma < 1$. The data generating processes stem from the time-varying Pareto-type model from Section 2.1 in the paper, that is,

$$P(Z_t > z) = z^{-1/\gamma} L_t(z).$$

Specifically, $T = 500$ observations over $\{t_j \equiv j/T\}_{j=1}^{T}$ are simulated from:

- **Scenario D**: $L_t(z) = t$ and $\gamma = 0.3$.
- **Scenario E**: $L_t(z) = \sqrt{t}$ and $\gamma = 0.5$.

Since these are scenarios of asymptotic independence the key parameter of interest is the coefficient of tail dependence. Again, we proceed as in Poon et al. (2003, 2004), i.e., only if there is no significant evidence to reject $\gamma = 1$ would we compute $f(t)$; otherwise, the processes is inferred to be asymptotically independent, and depending on the level γ they will be positively or negatively associated at the extremes. The Monte Carlo posterior median estimates of the coefficient of tail dependence are reported in Table 2 and provide an overall good fit.

To supplement these numerical experiments we also computed the EDIs for these scenarios. While the EDI is mostly tailored for AD it may also be computed under AI in some cases, such as in Scenarios D–E. For example, consider the Hall class of slowly-varying functions (Hall, 1982),

$$L_t(z) = c_0(t) + c_1(t) z^{-\beta(t)} + o(z^{-\beta(t)}), \tag{2}$$

where $c_0(t) > 0$, $\beta(t) > 0$, and $o(z^{-\beta(t)})$ is a remainder term. Under (2), it can be shown that $f(t) = c_0(t)/\int_0^1 c_0(\tau)\,\mathrm{d}\tau$. Scenarios D–E are particular cases of the Hall class with $c_0(t) = t$ and $c_0(t) = \sqrt{t}$, respectively. The posterior median EDI is presented in Fig. 2 and closely follows the target.

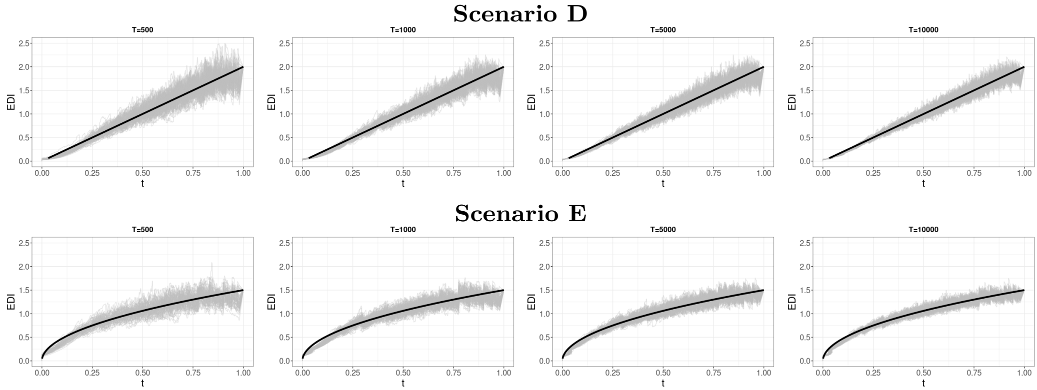

Fig. 2. Monte Carlo simulation study under asymptotic independence ($T = 500$): 150 randomly selected trajectories of EDI density estimates obtained via the posterior median of a mixture of finite Polya trees over the Monte Carlo simulation study (gray) plotted against true (black).

1.3. Multiwise EDI

The goal of this Monte Carlo simulation study is to assesses how the performance of the Polya tree-based EDI varies with the dimension of the multivariate vector. We extend Scenario A from the main paper by considering a time-varying multivariate logistic model

$$G_t(y_1, \dots, y_d) = \exp\left\{-\ell_t\left(\frac{1}{y_1}, \dots, \frac{1}{y_d}\right)\right\},$$

for $y_1, \dots, y_d > 0$, with

$$\ell_t(y_1, \dots, y_d) = (y_1^{1/\alpha_t} + \cdots + y_d^{1/\alpha_t})^{\alpha_t},$$

for $y_1, \dots, y_d > 0$, where $0 < \alpha_t \leq 1$. The true EDI can be derived using Di Bernardino and Rullière (2016, Theorem 2.2), and is given by

$$f(t) = \frac{-\sum_{i=1}^{d}(-1)^i \binom{d}{i} i^{\alpha_t}}{\int_0^1 \{-\sum_{i=1}^{d}(-1)^i \binom{d}{i} i^{\alpha_\tau}\} \, \mathrm{d}\tau}, \tag{3}$$

where $\binom{d}{i}$ is the binomial coefficient. Data $(\{Y_{1,t}, \dots, Y_{d,t}\})$ are simulated from a multivariate logistic extreme value copula with $\alpha_t = \sin(\pi t/T)$. Similar MCMC settings and prior information as in the paper are considered and we set u to be the 0.95 quantile of $\min\{Y_{1,t}, \dots, Y_{d,t}\}$. Table 3 shows the performance of the Polya tree-based EDI fits over different dimensions. As expected, for a fixed sample size, the accuracy of the fits is higher in the bivariate case.

1.4. Sensitivity analysis

In this section we conduct a prior sensitivity analysis via a Monte Carlo simulation study; the main results are tantamount to the ones obtained in Section 5 of the paper. We consider once again Scenarios A–C with $T = 1\,000$, where the baseline distribution is $F_{0,\theta}(t) = \beta(t; a, b)$, with $a \sim \text{Log-normal}(m_0, s_0)$, $b \sim \text{Log-normal}(\tau_1, \tau_2)$; for the precision parameter, we set $\alpha \sim \text{Gamma}(a_0, b_0)$. In terms of hyperparameters, we consider:

- Prior 1: $a_0 = .1$, $b_0 = .1$, $m_0 = .5$, $s_0 = 1$, $\tau_1 = .01$, and $\tau_2 = .01$.
- Prior 2: $a_0 = 1$, $b_0 = 1$, $m_0 = 1$, $s_0 = 1$, $\tau_1 = .01$, and $\tau_2 = .01$.

Table 3. Monte Carlo mean MISE (Mean Integrated Squared Error) for Scenarios A–C for Polya tree-based EDI.

Scenario A

Sample size (T)	Dimension (d) 2	3	4	5
500	0.0207	0.0781	0.1931	0.2187
1 000	0.0182	0.0589	0.1021	0.1597
5 000	0.0120	0.0519	0.0974	0.1252
10 000	0.0103	0.0496	0.0912	0.1150

Scenario B

Sample size (T)	Dimension (d) 2	3	4	5
500	0.0152	0.0437	0.0916	0.1711
1 000	0.0106	0.0309	0.0747	0.1084
5 000	0.0070	0.0159	0.0694	0.0901
10 000	0.0057	0.0063	0.0095	0.0193

Scenario C

Sample size (T)	Dimension (d) 2	3	4	5
500	0.0265	0.0546	0.0806	0.1467
1 000	0.0126	0.0192	0.0694	0.0946
5 000	0.0089	0.0125	0.0738	0.0974
10 000	0.0080	0.0091	0.0149	0.0237

Table 4. Test for the null hypothesis of a constant coefficient of tail dependence.

Pair of FAANG stocks	p-value
Facebook–Amazon	0.6064
Facebook–Apple	0.9084
Facebook–Netflix	0.7291
Facebook–Google	0.6930
Amazon–Apple	0.9979
Amazon–Netflix	0.4471
Amazon–Google	0.8669
Apple–Netflix	0.8148
Apple–Google	0.9661
Netflix–Google	0.3972

- Prior 3: $a_0 = .1$, $b_0 = .1$, $m_0 = \hat{\mu}$, $s_0 = \hat{\sigma}$, $\tau_1 = .01$, and $\tau_2 = .01$,

where $(\hat{\mu}, \hat{\sigma})$ is the maximum likelihood estimator of $\alpha \sim \text{Gamma}(a_0, b_0)$. For the threshold, we set once more u as the 0.95 quantile, and run a burn-in period of 5 000 iterates, after which we saved 5 000 posterior iterates. Fig. 3 reports the main findings of this Monte Carlo simulation study.

2. Supplementary empirical results

This section presents some supplementary empirical results complementing Section 5 of the paper. Section 2.1 provides the same empirical analysis as in the paper but allowing for the margins to be nonstationary. Section 2.4 summarizes results on goodness of fit, whereas Section 2.5 discusses EDI forecasts. Finally, Table 4 suggests that according to test T_4 in Einmahl et al. (2016, Corollary 2) there is no evidence to reject the null hypothesis of a constant coefficient of tail dependence.

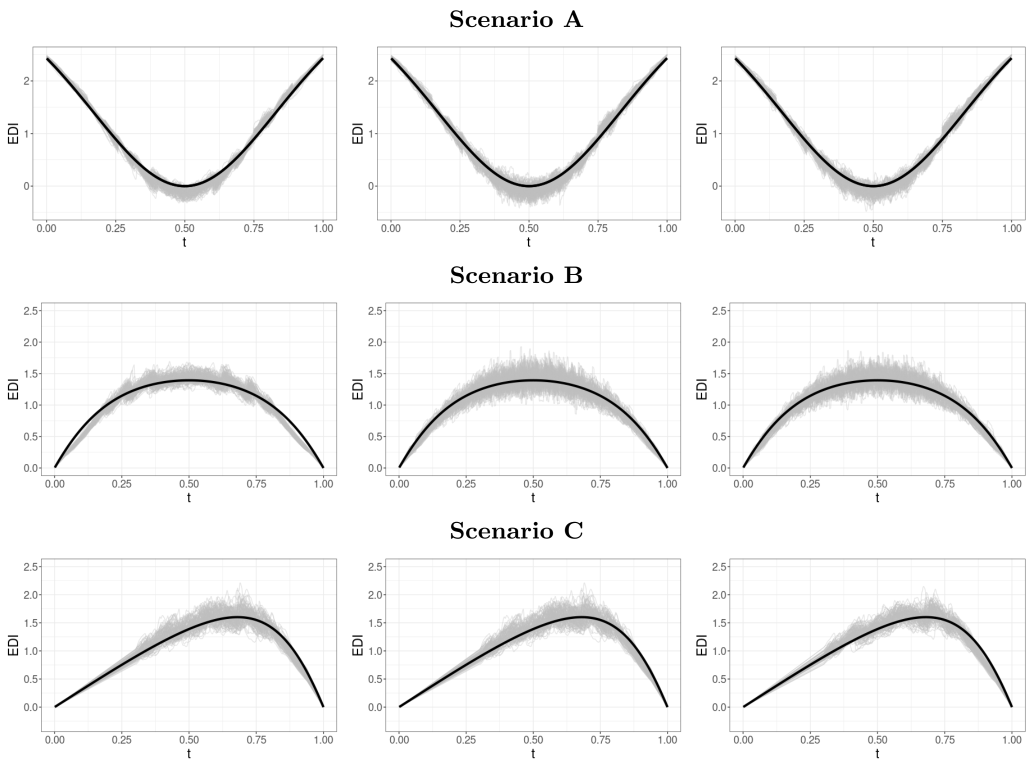

Fig. 3. Monte Carlo simulation study for sensitivity analysis ($T = 1\,000$): 150 randomly selected posterior median EDI density estimates, resulting from the the Monte Carlo simulation study, obtained via a mixture of finite Polya trees (gray), compared with the true EDI (black). From left to right: Priors 1, 2, and 3.

Table 5. Coefficient of tail dependence for FAANG stocks: Posterior median and 95% credible intervals for pairwise analysis.

Pair of FAANG stocks	Lower limit	Posterior median	Upper limit
Facebook–Amazon	0.982	1.000	1.000
Facebook–Apple	0.990	1.000	1.000
Facebook–Netflix	1.000	1.000	1.000
Facebook–Google	1.000	1.000	1.000
Amazon–Apple	1.000	1.000	1.000
Amazon–Netflix	0.968	1.000	1.000
Amazon–Google	0.988	1.000	1.000
Apple–Netflix	0.971	1.000	1.000
Apple–Google	0.998	1.000	1.000
Netflix–Google	0.964	1.000	1.000

2.1. Nonstationary margins

We transform the bivariate returns (R_t^X, R_t^Y) to unit Fréchet margins (X_t, Y_t) using the transformation:

$$(X_t, Y_t) = (-1/\log G_t(R_t^X), -1/\log H_t(R_t^Y)), \tag{4}$$

where G_t and H_t are the respective marginal time-varying distribution functions for R_t^X and R_t^Y. To transform the data as in (4) we compute the time-varying distribution function estimator of Harvey and Oryshchenko (2012). That is

$$\widehat{G}_t(x) = \sum_{i=1}^{T} K\left(\frac{x - R_i^X}{h}\right) w_{t,i}, \quad t = 1, \dots, T,$$

where $K(\cdot)$ is a kernel distribution function, and $\sum_{i=1}^{T} w_{t,i} = 1$ for all t. We use the Gaussian kernel and the weights are computed using the algorithm of Koopman and Harvey (2003), where it follows that $w_{t,i} \approx (1-w)/(1+w)w^{|t-i|}$ for $i = 1, \dots, T$ and $0 \leq w < 1$. We tried different values of the parameter w and obtained similar findings; the results reported here are those for $w = 0.6$. Trivially, $\widehat{H}_t$ is analogously defined. Fig. 4 shows the pairwise EDI and Fig. 5 depicts the multiwise EDI. As can be seen from the latter charts, the key empirical findings are the same as those reported in Section 5 of the paper. Table 2.1 presents the posterior median and credible bands of the coefficient of tail dependence. The posterior median coefficient of tail dependence for this multivariate analysis is 0.68 (CI = $(0.60, 0.83)$). Again, the results are tantamount to the ones obtained in the paper.

2.2. GARCH filtering

Fig. 6 shows the pairwise EDI's after prewhitening the returns using an asymmetric version of the GARCH (AGARCH) model, following Poon et al. (2003, Appendix A.2). Fig. 9 depicts the multiwise EDI after prewhitening the returns using the AGARCH approach. The main empirical findings on the key dynamics of extremal dependence are largely consistent with those reported in the main paper, though as expected the resulting EDIs show slight variations.

2.3. Subperiod analysis

This section comments on some links between the EDI and the subperiod estimator of Poon et al. (2003, Section 3.3.2) for $\chi(t)$. Specifically, we show that an histogram-type of estimator

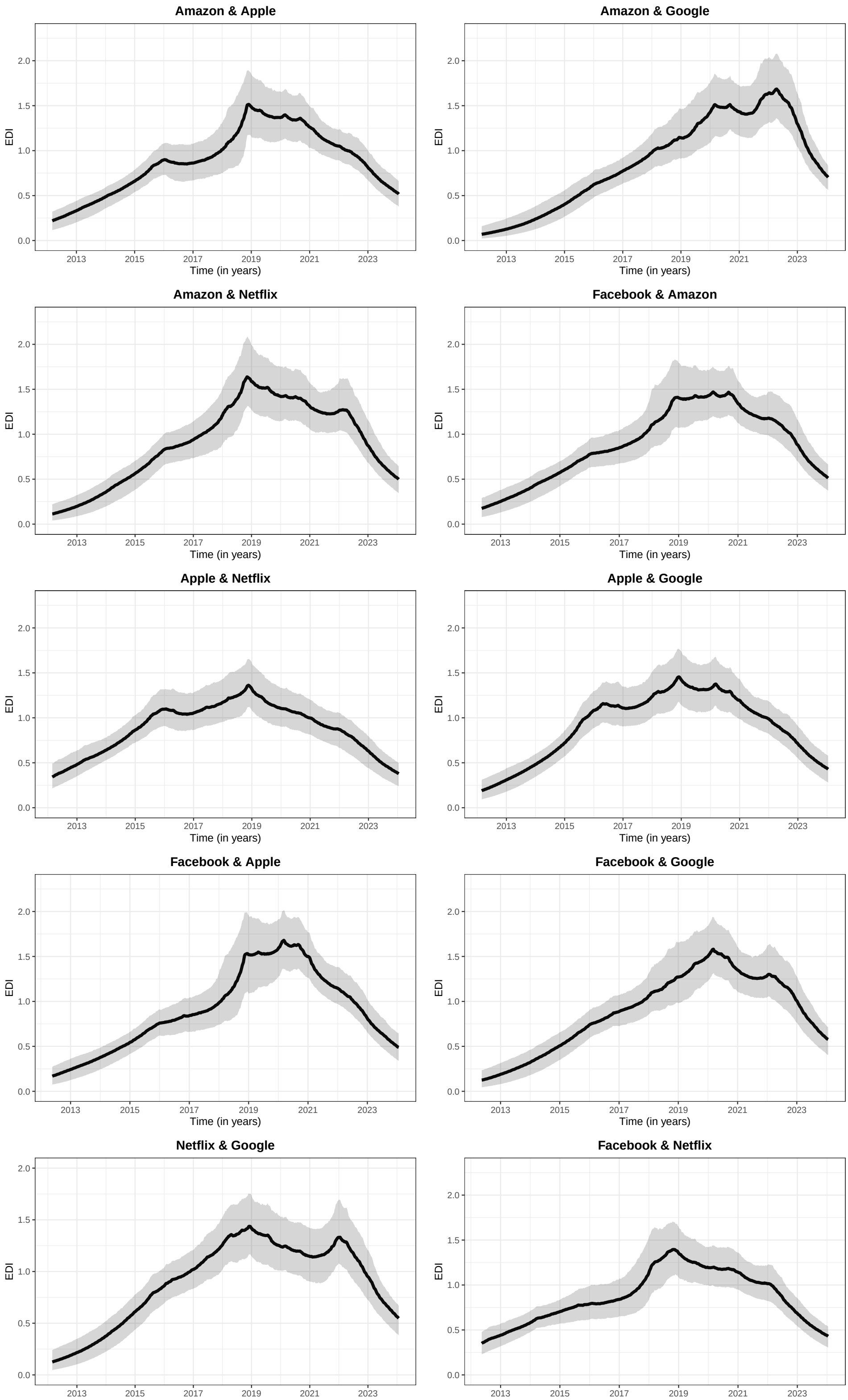

Fig. 4. Pairwise EDI for FAANG stocks: Posterior median of EDI based on a mixture of finite Polya trees along with pointwise credible bands.

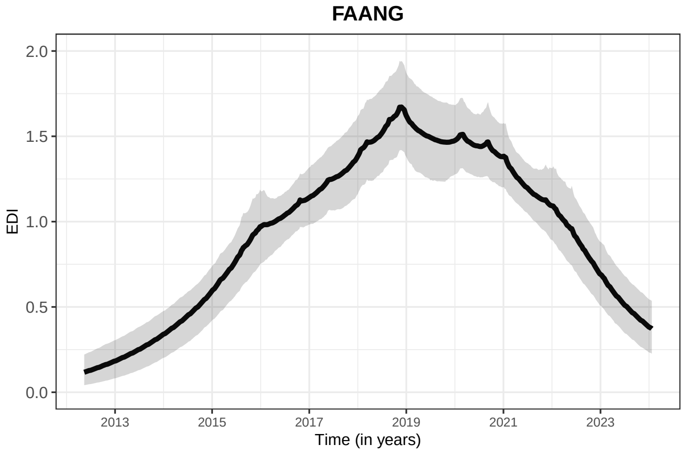

Fig. 5. Multiwise EDI for FAANG stocks: Posterior median EDI based on a mixture of finite Polya trees along with pointwise credible bands.

for the EDI is related with the subperiod estimator. To show this, let $I = \{\tau_1, \dots, \tau_k\}$ be the times of the exceedances and partition the unit interval $(0, 1)$ into m bins,

$$B_1 = \left(0, \frac{1}{m}\right), \dots, B_m = \left(\frac{m-1}{m}, 1\right).$$

The subperiod estimator of Poon et al. can be written as

$$\chi(t) = \sum_{j=1}^{m} \hat{\chi}_j \mathbb{1}_{\{t \in B_j\}} = \sum_{j=1}^{m} \frac{un_u^j}{n_j} \mathbb{1}_{\{t \in B_j\}} = \sum_{j=1}^{m} \frac{n_u^j}{n_j/u} \mathbb{1}_{\{t \in B_j\}}. \tag{5}$$

Here, $\hat{\chi}_j = (un_u^j)/n_j$, $n_u^j = \sum_{t \in I} \mathbb{1}_{\{t \in B_j\}}$, and $n_j = \sum_{t \in \mathbb{T}} \mathbb{1}_{\{t \in B_j\}}$ with $j = 1, \dots, m$.

Yet, (5) is not a density estimator and hence it is not a fair comparison to the EDI, as by definition $f(t) = \chi(t)/\int_0^1 \chi(t)\,\mathrm{d}t$. Still, as shown below, the estimator of $\chi(t)$ in (5) has some links with the histogram estimator of the EDI, which is defined as

$$f(t) = \sum_{j=1}^{m} \frac{n_u^j}{k/m} \mathbb{1}_{\{t \in B_j\}}. \tag{6}$$

This is similar to the subperiod estimator, but the need for $f(t)$ to integrate to one is what justifies the different denominators when comparing (5) against (6). Fig. 8 shows the histogram estimator of the pairwise EDI, along with the proposed mixture of finite Polya trees estimator; Sturge's rule was used to set the number of bins. As can be seen from Fig. 8 the fitted histogram EDIs are in line with those reported in the paper.

2.4. *QQ-plots from randomized quantile residuals*

To assess the fit of the proposed methods we resort to a version of randomized quantile residuals (Dunn and Smyth, 1996), where residuals are defined as $\varepsilon_j = \Phi^{-1}\{F(\tau_j)\}$, for $j = 1, \dots, k$; here, F is the integrated EDI function and τ_j the standardized time of the exceedances. The rationale for such residuals is that if F is the true distribution of the time of the exceedances, then $\tau_j \mid F \overset{\text{iid}}{\sim} F$, implying that $F(\tau_j)$ should be Uniform, and thus ε_j

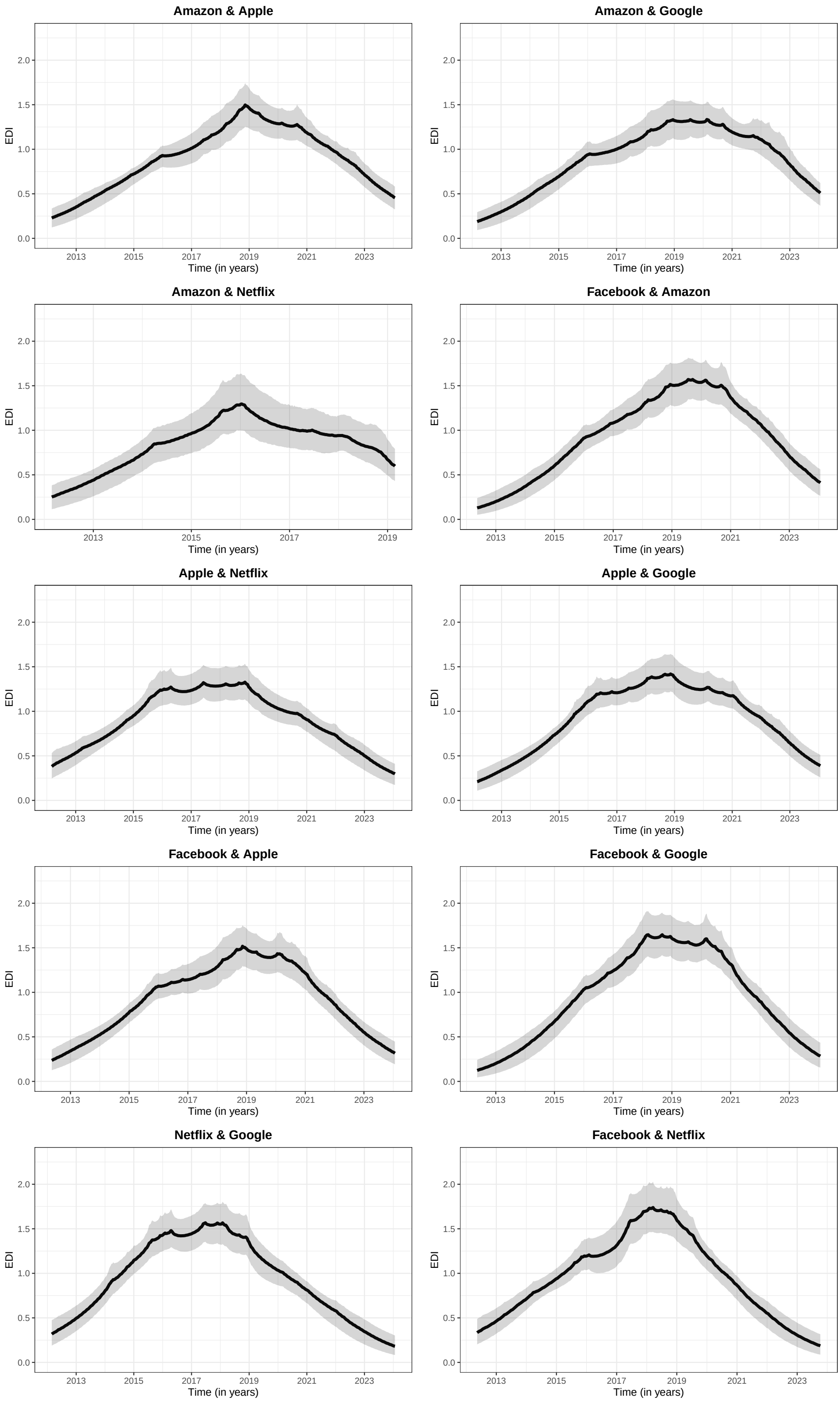

Fig. 6. Pairwise EDI for FAANG stocks: Posterior median of EDI based on a mixture of finite Polya trees along with pointwise credible bands, when filtering the margins using an AGARCH.

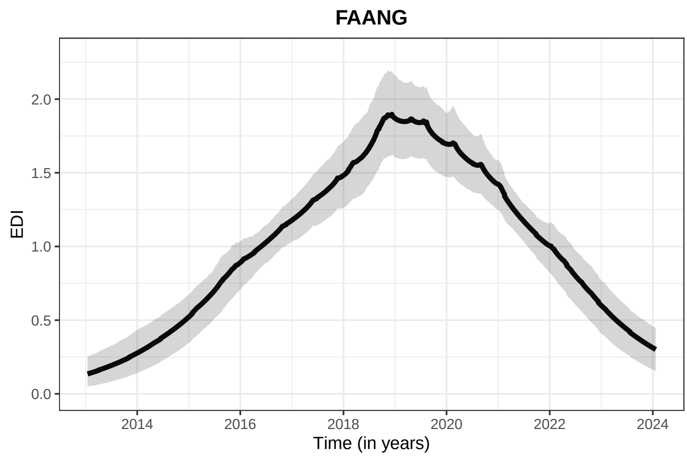

Fig. 7. Multiwise EDI for FAANG stocks: Posterior median EDI based on a mixture of finite Polya trees along with pointwise credible band, when filtering the margins using an AGARCH.

should be Normal distributed, for all j. Figs. 10–11 depict QQ-plots of randomized quantile residuals plotted against the theoretical standard Normal quantiles, and suggest acceptably good fits of the proposed model—both in the Monte Carlo simulation study from Section 5 and in the real data analysis from Section 6.

2.5. A discretized Holt–Winters forecasting approach for the EDI function

The future evolution of tail-dependence structure can be predicted by treating each MCMC trajectory of the EDI as a stochastic process to be forecasted. Specifically, we consider an Holt–Winters additive method based on three smoothing equations (level, trend, and seasonality). Related approaches can be found in Hyndman et al. (2008, Section 2.3.4); in principle, other forecasting methods for continuous time processes (Harvey, 1990, Chapter 9) can also be applied with the due modifications.

Specifically, let $\{\ell(t)\}_{t\in\mathbb{T}_0} = \{\log f(t)\}_{t\in\mathbb{T}_0}$ be a log posterior sampled EDI over an equally-spaced grid $\mathbb{T}_0 = \{1/T, \dots, (T-1)/T, 1\}$. Before obtaining the resulting EDI, we first elongate the EDI outside the unit interval; then, later below we will map back time back to the unit interval. To elongate the EDI to $[1, 1 + h/T]$, for any $h > 0$, we consider following Holter–Winter specification,

$$\exp\{\tilde{\ell}(1 + h/T)\} = l(1) + b(1)h/T + s(1 - p + h_p^+/T), \tag{7}$$

where p is the length of seasonality, and the smoothing equations are

$$\begin{cases} l(1) & = \alpha\{f(1) - s(1-p)\} + (1-\alpha)\{l(1-1/T) + b(1-1/T)\}, \\ b(1) & = \beta\{l(1) - l(1-1/T)\} + (1-\beta)b(1-1/T), \\ s(1) & = \gamma\{f(1) - l(1-1/T) - b(1-1/T)\} + (1-\gamma)s(1-p). \end{cases}$$

Here, $h_p^+ = \{(h-1) \mod p\} + 1$ and $0 < \alpha, \beta, \gamma < 1$ are smoothing constants, one for each of the respective smoothing equations. The elongation obtained in (7) will now be blended with the original EDI, $f(t)$. Hence, the EDI resulting from this will include the forecast for

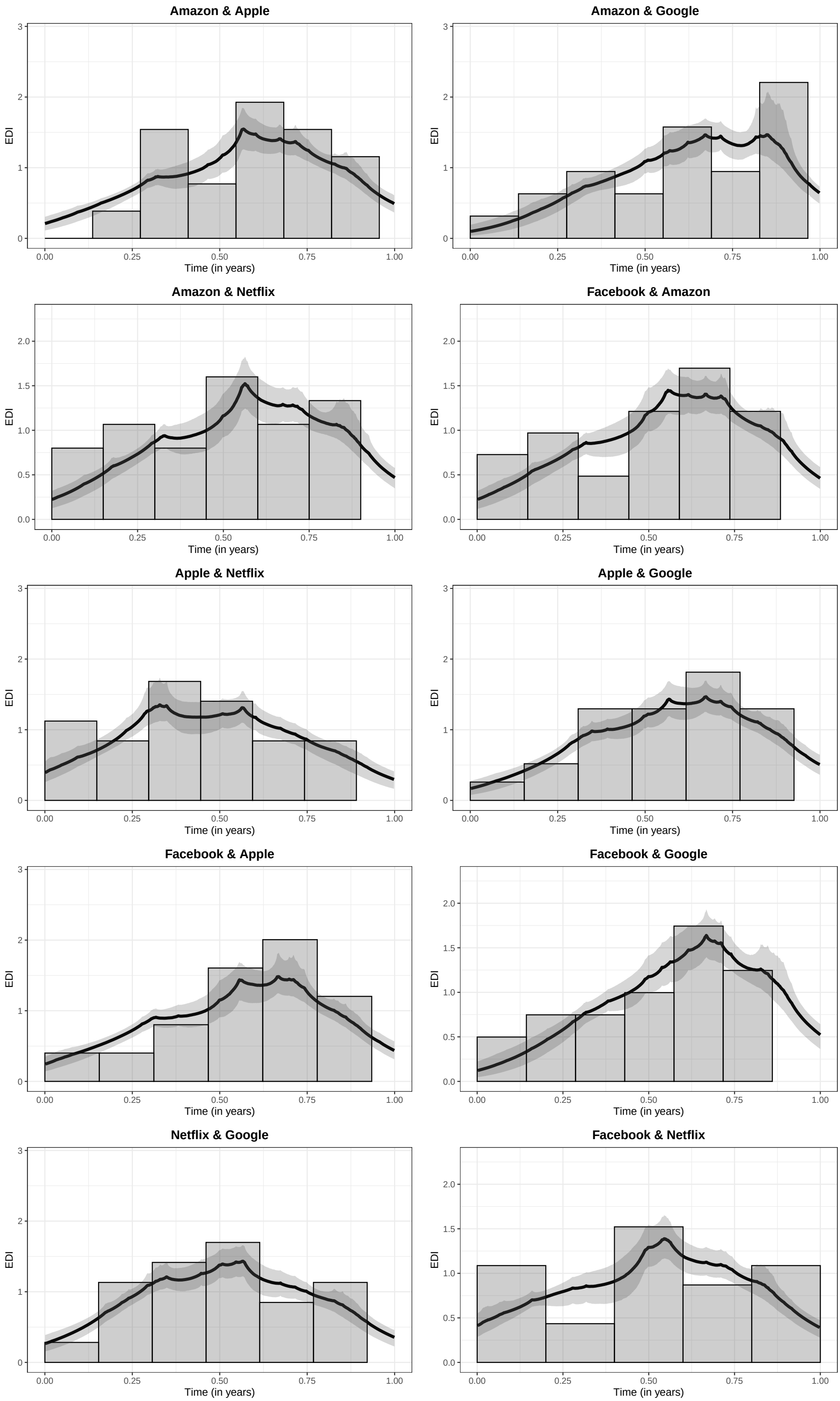

Fig. 8. Pairwise EDI for FAANG stocks: Posterior median of EDI based on a mixture of finite Polya trees with pointwise credible bands, along with histogram estimator of the pairwise EDI.

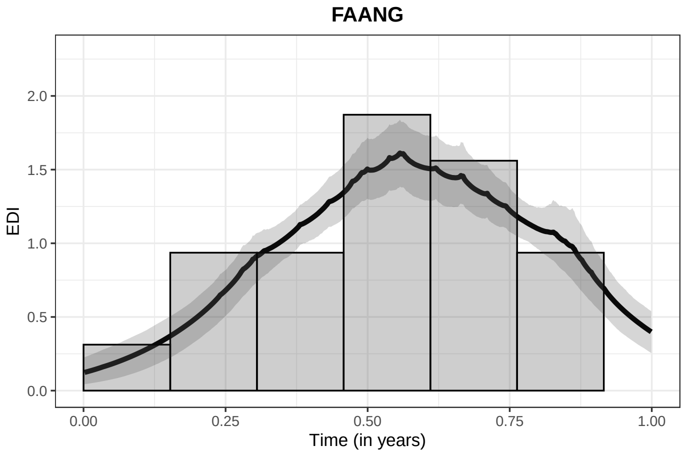

Fig. 9. Multiwise EDI for FAANG stocks: Posterior median EDI based on a mixture of finite Polya trees with pointwise credible bands, along with histogram estimator of the pairwise EDI.

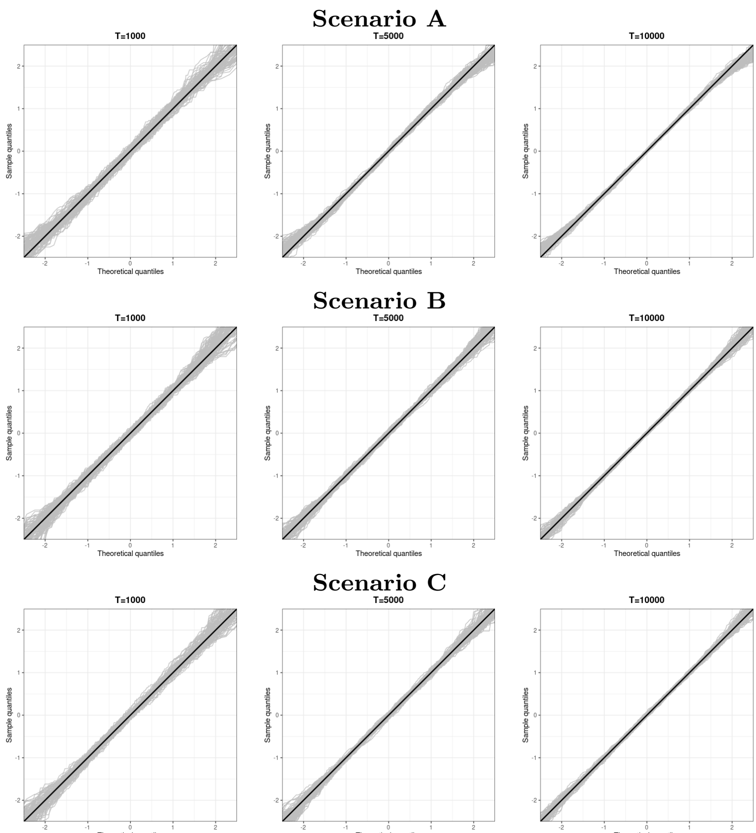

Fig. 10. Goodness of EDI fits for Monte Carlo simulation: QQ-plots of randomized quantile residuals for 150 randomly selected trajectories of EDI density estimates obtained via the posterior median of a mixture of finite Polya trees over the Monte Carlo simulation study from Section 4 in the paper.

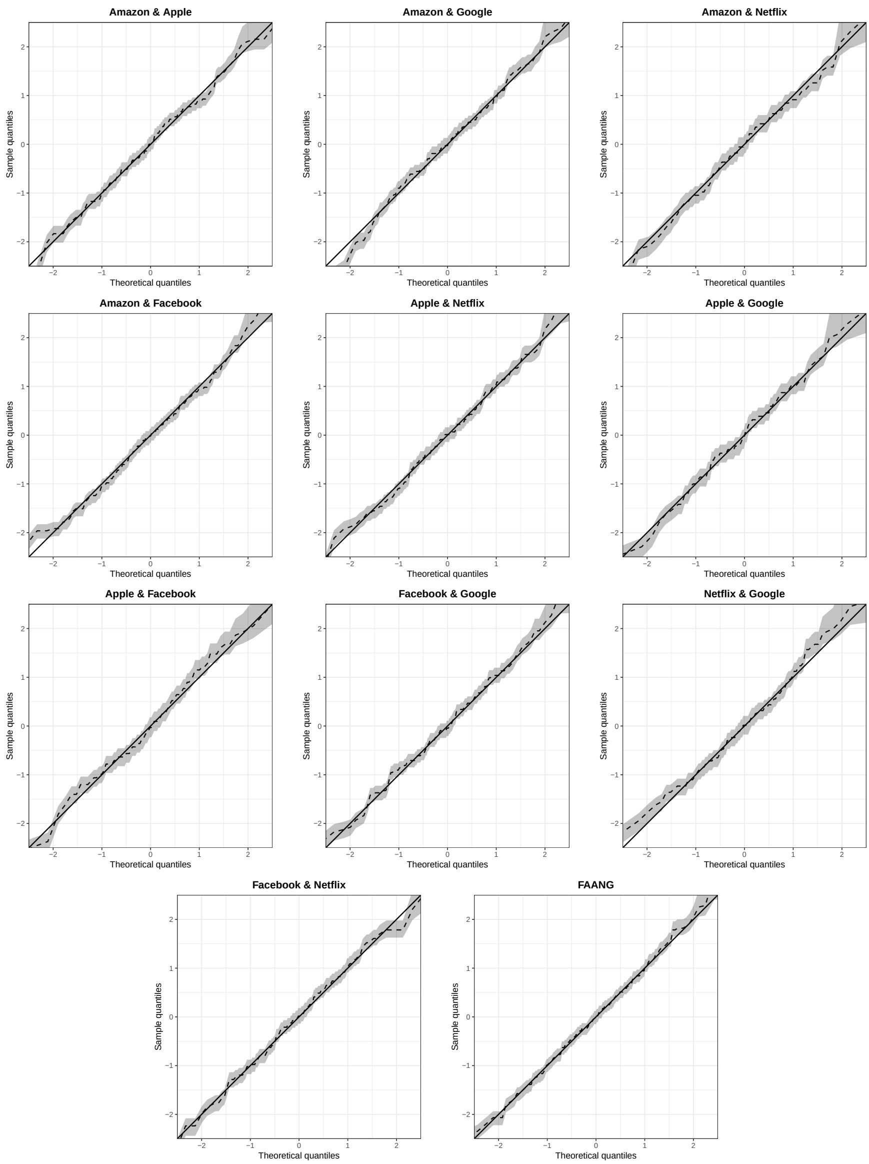

Fig. 11. Goodness of EDI fits for FAANG stocks: QQ-plots of randomized quantile residuals (for pairwise and multiwise analyses) of fitted mixture of finite Polya trees for FAANG stocks over 2012–2024.

h periods ahead, and is given by

$$\tilde{f}\left(\frac{i}{T+h}\right) \propto \begin{cases} f(i/T), & i \in \{1, \ldots, T\}, \\ \exp\{\tilde{\ell}(i/T)\}, & i \in \{T+1, \ldots, T+h\}. \end{cases} \tag{8}$$

That is, from (8), we derive an EDI that visually corresponds to $f(t)$ across the first T points on a newly rescaled time grid; beyond that point, it is determined by the Holter–Winter prediction. Formally, the resulting $\tilde{f}(t)$ is evaluated over the grid,

$$\mathbb{T}_1 = \Bigg\{ \underbrace{\frac{1}{T+h}, \ldots, \frac{T}{T+h}}_{\text{adjusted times for observed}}, \underbrace{\frac{T+1}{T+h}, \ldots, 1}_{h \text{ forecasts}} \Bigg\}.$$

The construction of EDI forecasts above has taken into account the need for incorporating a positivity and a normalization constraint (i.e., $\tilde{f}(t) > 0$ and $\int_0^1 \tilde{f}(t)\,\mathrm{d}t = 1$). Trivially it follows from (8) that $\tilde{f}(t) > 0$, while the normalizing constant in (8) can be made precise and is given by

$$1 + \int_1^{1+h/T} \exp\{\tilde{\ell}(1+u)\}\,\mathrm{d}u = \frac{1}{T}\left[\sum_{i=1}^{T} f(i/T) + \sum_{i=T+1}^{T+h} \exp\{\tilde{\ell}(i/T)\}\right] + o(1),$$

as $T \to \infty$.

Fig. 12 shows the resulting EDI forecast; following standard practices in the literature (e.g., Hyndman et al., 2008), α, β and γ are set to minimize the squared one-step prediction error. The fits are illustrated in Fig. 12; for instance, the EDI forecast for Amazon–Apple suggests a potential increase in the frequency of joint extreme losses for this pair in 2025. Conversely, as indicated in the same figure, a decrease in the frequency of joint extreme losses is anticipated over the same period for other pairs (e.g., Facebook–Google).

References

Di Bernardino, E. and Rullière, D. (2016), On tail dependence coefficients of transformed multivariate Archimedean copulas, *Fuzzy Sets and Systems*, 284, 89–112.

Dunn, P. K. and Smyth, G. K. (1996), Randomized quantile residuals, *Journal of Computational and Graphical Statistics*, 5, 236–244.

Einmahl, J. H. J., de Haan, L., and Zhou, C. (2016), Statistics of heteroscedastic extremes, *Journal of the Royal Statistical Society: Series B*, 78, 31–51.

Hall, P. (1982), On some simple estimates of an exponent of regular variation, *Journal of the Royal Statistical Society: Series B (Methodological)*, 44, 37–42.

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

Harvey, A. C. (1990), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, MA: Cambridge University Press.

Hyndman, R., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008), *Forecasting with Exponential Smoothing: The State Space Approach*, New York: Springer.

Koopman, S. J. and Harvey, A. (2003), Computing observation weights for signal extraction and filtering, *Journal of Economic Dynamics and Control*, 27, 1317–1333.

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.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, London: CRC.

Wand, M. P. and Jones, M. C. (1995), *Kernel Smoothing*, London: Chapman & Hall.

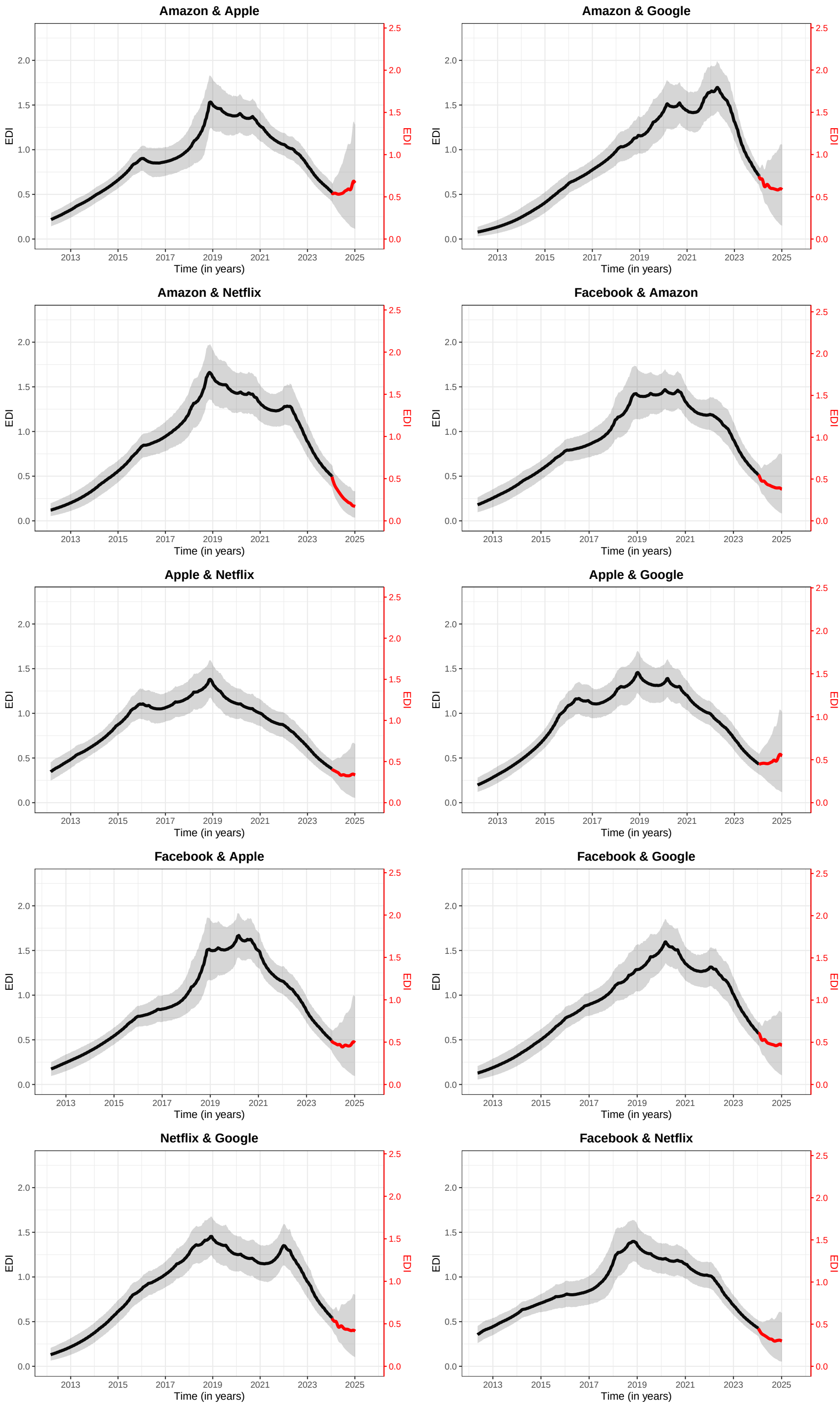

Fig. 12. Pairwise EDI for FAANG stocks: EDI Holt–Winters forecast with pointwise credible bands.