Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: https://www.tandfonline.com/loi/ubes20

Dynamic Bivariate Peak over Threshold Model for Joint Tail Risk Dynamics of Financial Markets*

Zifeng Zhao

To cite this article: Zifeng Zhao (2020): Dynamic Bivariate Peak over Threshold Model for Joint Tail Risk Dynamics of Financial Markets*, Journal of Business & Economic Statistics, DOI: 10.1080/07350015.2020.1737083

To link to this article: https://doi.org/10.1080/07350015.2020.1737083

Dynamic Bivariate Peak over Threshold Model for Joint Tail Risk Dynamics of Financial Markets*

Zifeng Zhao

Mendoza College of Business, University of Notre Dame

*We thank the Editor, the Associate Editor and two referees for their comments and suggestions that have led to a substantial improvement in the quality of this paper.

Corresponding author Zifeng Zhao zzhao2@nd.edu

Abstract

We propose a novel dynamic bivariate peak over threshold (PoT) model to study the time-varying behavior of joint tail risk in financial markets. The proposed framework provides simultaneous modeling for dynamics of marginal and joint tail risk, and generalizes the existing tail risk literature from univariate dimension to multivariate dimension. We introduce a natural and interpretable tail connectedness measure and examine the dynamics of joint tail behavior of global stock markets: empirical evidence suggests markets from the same continent have time-varying and high-level joint tail risk, and tail connectedness increases during periods of crisis. We further enrich the tail risk literature by developing a novel portfolio optimization procedure based on bivariate joint tail risk minimization, which gives promising risk-rewarding performance in backtesting.

Keywords: time-varying tail risk, multivariate extreme value theory; dynamic modeling; financial risk management; portfolio optimization; tail connectedness

1 Introduction

The study of extreme events (i.e. large losses) is one of the central topics in financial risk management, which concerns with the tail behavior of financial markets. An accurate assessment of tail behavior among financial assets is essential for risk minimization, portfolio optimization, and for monitoring stability

of financial markets. One of the most valuable tools for studying tail behavior of financial markets is Extreme Value Theory (EVT), which is proven to be successful in predicting conditional Value-at-Risk (Chavez-Demoulin et al., 2005; McNeil and Frey, 2000), in tracking dynamics of tail risk (Massacci, 2016; Kelly and Jiang, 2014), and in understanding the behavior of financial markets during crisis (Longin and Solnik, 2001; Quintos et al., 2001; Poon et al., 2004). The behavior of financial assets is reflected in the time series of its losses (i.e. negative returns), motivating the vast literature in financial time series modeling.

A stylized fact of financial time series is its time-varying behavior, where certain properties of the series exhibit dynamics across time. The most notable example is volatility clustering, where the conditional variance changes over time with large values clustering together. See for example the well-known ARCH (Engle, 1982) and GARCH (Bollerslev, 1986) literature. Recently, there has been an increased interest in studying the dynamics of higher order behavior, especially the tail behavior, of financial time series. Utilizing univariate EVT-based time series models, Kelly and Jiang (2014), Massacci (2016), Mao and Zhang (2018), and Zhao et al. (2018) find substantial evidence that the tail index of the stock market exhibits significant dynamics over time, after controlling the effects of volatility clustering. In other words, the marginal tail behavior of individual financial assets tends to be time-varying.

For financial risk management of a portfolio containing multiple assets, besides marginal behavior of each individual asset, it is even more crucial to understand the joint behavior of multiple assets, especially their joint *tail* behavior, which concerns with the relationship of multiple financial assets after their losses exceed certain high thresholds. Unlike the aforementioned univariate case, multivariate EVT-based methodology for studying dynamics of joint tail behavior is scarce in financial econometrics and risk management literature, making the

development of a dynamic model for tracking joint tail risk a compelling and critical matter.

For multivariate financial time series, most of the existing research focuses on the dynamics of the overall dependence structure, such as correlation, for example see DCC in Engle (2002) and Pakel et al. (2018). However, correlation structure cannot help reveal the joint tail behavior as it describes the overall relationship between variables and only captures linear dependence. More recently, copula has gained popularity in constructing models for dynamic dependence of multivariate time series, thanks to its ability to capture non-linear and tail dependence, see Creal et al. (2011), Eckernkemper (2017) and Oh and Patton (2018). However, the existing literature specifies the dynamic copula in a fully parametric fashion, which may not be robust against model misspecification. Moreover, same as the dynamic correlation literature, the dynamic copula literature aims to model the overall dependence instead of focusing on the joint tail behavior. As is discussed in the EVT literature (e.g. Poon et al., 2004; Bosma et al., 2019), when the primary interest is the tail behavior, observations from the non-extremal part of the distribution may contaminate the analysis.

Motivated by the above discussions, in this paper we aim to develop a dynamic model that is designed exclusively for capturing the dependence of extremal observations of multivariate time series, which thus provides a new tool for rigorous studies of joint tail behavior in financial markets. Given a d-variate time series $\{\mathbf{X}_t = (X_{t1}, \ldots, X_{td})\}_{t=1}^{T}$, the extremal observations for each univariate time series $\{X_{ti}\}_{t=1}^{T}$ can be defined as $\{Y_{ti} = \max(X_{ti} - \tau_i, 0)\}_{t=1}^{T}$, where τ_i denotes a high threshold. A popular choice in the EVT literature is to take τ_i as an extreme quantile, e.g. 90% or 95% quantile, of the observed time series $\{X_{ti}\}_{t=1}^{T}$. Thus, the censored time series $\{\mathbf{Y}_t = (Y_{t1}, \ldots, Y_{td})\}_{t=1}^{T}$ contains information solely about joint tail behavior of $\mathbf{X}_t$, where some component time series X_{ti} observes an extreme outcome.

We focus on the censored time series $\{\mathbf{Y}_t\}$ and develop a novel semiparametric time series model for dynamic behavior of $\{\mathbf{Y}_t\}$, which in turn reveals joint tail behavior of $\{\mathbf{X}_t\}$. Specifically, generalizing the classical multivariate EVT to time series context, we propose a dynamic multivariate Peak-over-Threshold (PoT) model. This approach avoids fully parametric assumptions on $\mathbf{X}_t$ and minimizes the impact of non-extremal behavior of $\mathbf{X}_t$ (i.e. when $X_{ti} < \tau_i, i = 1, \ldots, d$) on the inference, and thus provides an accurate and rigorous statistical framework to study the dynamics of joint tail behavior of financial markets.

This paper provides a threefold contribution to the growing literature on tail risk measurement in financial markets. First and foremost, we generalize the existing literature on tail risk from univariate/marginal dimension to multivariate/joint dimension by proposing the dynamic multivariate PoT model. It offers a novel EVT-based econometric tool to quantify the time-varying joint tail behavior of financial markets and can be seen as a natural generalization of the existing dynamic models for marginal tail behavior of univariate time series (Massacci, 2016; Zhang and Schwaab, 2017). In addition, a more interpretable and sensible dynamic model for marginal tail risk is derived in the development of the dynamic multivariate PoT model. Second, in the empirical study, based on the proposed model, we introduce a natural measure of tail connectedness between different financial assets and further examine the dynamics of joint tail risk of global financial markets. We find that markets within the same continent tend to have high-level and time-varying joint tail risk and markets from different continents tend to exhibit relatively low-level and stable joint tail risk. Moreover, tail connectedness among global stock markets increases during crisis. Third, a novel portfolio optimization procedure is further proposed based on the dynamic PoT model and is demonstrated to generate promising risk-rewarding portfolios under both normal state and crisis state markets. This helps expand the potential applications of tail risk literature to the area of active financial risk management.

The rest of the paper is organized as follows. In Section 2, we derive the dynamic multivariate PoT model and its statistical properties. The estimation and prediction procedures are presented in Section 3, along with a simulation study. Section 4 studies the dynamics of joint tail risk of global financial markets and Section 5 presents a novel portfolio optimization framework based on bivariate joint tail risk minimization. We conclude the paper in Section 6. The supplementary material contains technical materials and additional empirical studies.

2 Dynamic Multivariate Peak over Threshold Model

2.1 Preliminary

Recall that by definition, $Y_{ti} = \max(X_{ti} - \tau_i, 0)$ with τ_i being a high threshold. In other words, the censored observations $\mathbf{Y}_t = (Y_{t1}, \ldots, Y_{td})$ corresponds to the joint behavior of $\mathbf{X}_t$ over high threshold $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)$. Thus, to derive a sensible and interpretable model for $\mathbf{Y}_t$, we resort to the analysis of $\mathbf{X}_t$ by employing the classical multivariate EVT (Resnick, 1987; Diebold et al., 1998; Embrechts et al., 1999). For simplicity of presentation, in this paper we focus our discussion on the bivariate case with $d = 2$, though the ideas and techniques extend naturally to higher dimensions.

On day t, conditional on past observations $\mathcal{F}_{t-1} = \boldsymbol{\sigma}(\mathbf{X}_s, s < t)$, we denote the joint cdf (cumulative distribution function) of $\mathbf{X}_t = (X_{t1}, X_{t2})$ as $F_{\mathbf{X}_t}(x_1, x_2)$ and the marginal cdfs as $F_{X_{t1}}(x)$ and $F_{X_{t2}}(x)$. By Sklar's theorem (Sklar, 1959), there exists a copula $C_t^0(u_1, u_2)$ such that

$$F_{\mathbf{X}_t}(x_1, x_2) = C_t^0(F_{X_{t1}}(x_1), F_{X_{t2}}(x_2)),$$

which decomposes the behavior of $\mathbf{X}_t$ into marginal cdfs $F_{X_{ti}}(x), i=1,2,$ and a copula C_t^0. For more details about copula, we refer to Nelsen (1999) and Joe (2014).

We assume that the marginal cdfs $F_{X_{t1}}(x)$ and $F_{X_{t2}}(x)$ are heavy-tailed and are in the Domain of Attraction (DoA) of the Fréchet distribution (Leadbetter et al., 1983), which is a common assumption about financial time series. Thus, we have that

$$1-F_{X_{ti}}(x) \sim L_{ti}(x)x^{-\xi_{ti}}, \qquad (1)$$

where $L_{ti}(x)$ is a slowly varying function (Resnick, 1987) and $\xi_{ti}>0$ is called the tail index of X_{ti}. Here and after, $f(x) \sim g(x)$ means that $f(x)/g(x) \to 1$ as $x \to \infty$. Condition (1) is also referred to as the regular variation of the univariate cdf $F_{X_{ti}}(x)$. Note that (1) provides a semiparametric description of the marginal tail behavior of X_{ti} in the sense that $F_{X_{ti}}(x) \approx 1-L_{ti}(x)x^{-\xi_{ti}}$, for $x>\tau_i$ with some high threshold τ_i. In Section 2.2, based on (1), we propose a dynamic univariate PoT model for univariate time series $\{Y_{ti}\}_{t=1}^T$.

To characterize the behavior of copula C_t^0, we further assume that the joint cdf $F_{\mathbf{X}_t}(x_1,x_2)$ is in the DoA of a bivariate extreme value distribution (Resnick, 1987). Thus, there exists an extreme value copula $C_t(u_1,u_2)$, see Segers (2012), such that $F_{\mathbf{X}_t}(x_1,x_2)$ can be approximated in the sense that

$$F_{\mathbf{X}_t}(x_1,x_2) \approx C_t(F_{X_{t1}}(x_1),F_{X_{t2}}(x_2)), \qquad (2)$$

for $(x_1,x_2) \in \{(x_1,x_2) \mid x_1>\tau_1, x_2>\tau_2\}$ with high threshold $\boldsymbol{\tau}$. In other words, when (x_1, x_2) is in the joint tail region, we can find an extreme value copula C_t which provides an accurate approximation for the original copula C_t^0. The static version

of the distribution in (2) recovers the multivariate threshold model in Ledford and Tawn (1996).

As discussed above, $\mathbf{Y}_t$ corresponds to the behavior of $\mathbf{X}_t$ over high threshold $\boldsymbol{\tau}$. Thus, together, (1) and (2) provide a fundamental theoretical ground for the modeling of $\mathbf{Y}_t$, where the marginal behavior of Y_{ti} is characterized by the univariate regular variation of $F_{X_{ti}}$ in (1) and the joint behavior of (Y_{t1}, Y_{t2}) is captured by the extreme value copula C_t in (2). In the following two subsections, we use (1) and (2) as the basis for our proposed dynamic bivariate PoT model, where we further develop sensible and interpretable dynamics to model the possibly time-varying marginal and joint behavior of (Y_{t1}, Y_{t2}).

2.2 Dynamic univariate PoT model

In this section, we propose the dynamic univariate PoT (uPoT) model where dynamics is designed for the time-varying marginal tail behavior of univariate time series $\{Y_{ti}\}_{t=1}^{T}$. We build the dynamic uPoT model based on a more elaborate version of (1), which incorporates the time-varying dynamics of financial time series $\{X_{ti}\}_{t=1}^{T}$.

Following the well-established GARCH literature (Engle, 1982; Bollerslev, 1986), the observed time series $\{X_{ti}\}$ can be modeled as a scale family such that

$$X_{ti} = \sigma_{ti} e_{ti}, \quad (3)$$

where σ_{ti} is the time-varying scale parameter and $\{e_{ti}\}$ is assumed to be a sequence of independent random innovations with $E(e_{ti}) = 0$ and $\mathrm{Var}(e_{ti}) = 1$.

We derive the marginal distribution of Y_{ti} with (3) as the underlying time series model, however, we relax the unit-variance assumption of GARCH and only require the regular variation of e_{ti}. Specifically, we assume that e_{ti} exhibits a Pareto-type tail behavior such that,

$$P(e_{ti} > z) = 1 - F_{e_{ti}}(z) \sim k_i z^{-\xi_{ti}}, \qquad (4)$$

where $\xi_{ti} > 0$ is the time-varying tail index of e_{ti}. Note that (4) is a special case of (1) with the slowly varying function being a constant k_i. A broad class of random variables such as Cauchy, Lévy, Pareto, and *t*-distributions satisfy (4). For simplicity and model parsimony, without loss of generality, we assume $k_i = 1$ since the scale parameter σ_{ti} can help absorb the effect of k_i.

Based on (4), for a suitably chosen high threshold τ_i, we have that $P(Y_{ti} = 0) = P(X_{ti} \le \tau_i) = P(e_{ti} \le \tau_i / \sigma_{ti}) = 1 - (\tau_i / \sigma_{ti})^{-\xi_{ti}}$ and $P(Y_{ti} \le y \mid Y_{ti} > 0) = P(X_{ti} - \tau_i \le y \mid X_{ti} - \tau_i > 0) = P(e_{ti} \le (y + \tau_i) / \sigma_{ti} \mid e_{ti} > \tau_i / \sigma_{ti}) = 1 - (1 + y / \tau_i)^{-\xi_{ti}}$. The distribution function of Y_{ti} on day t is thus

$$\begin{aligned} F_{Y_{ti}}(y) &= 1 - \left(\tau_i / \sigma_{ti}\right)^{-\xi_{ti}} \left(1 + y / \tau_i\right)^{-\xi_{ti}}, \\ f_{Y_{ti}}(y) &= \left(1 - \left(\tau_i / \sigma_{ti}\right)^{-\xi_{ti}}\right) I(y = 0) + \xi_{ti} \sigma_{ti}^{\xi_{ti}} (\tau_i + y)^{-\xi_{ti} - 1} I(y > 0), \end{aligned} \qquad (5)$$

where $f_{Y_{ti}}(y)$ denotes the density function, and the marginal behavior of Y_{ti} is characterized by the scale parameter σ_{ti} and the tail index ξ_{ti}, which possibly vary across time t.

To design the time-varying dynamics for (σ_{ti}, ξ_{ti}), one possibility is to utilize the popular generalized autoregressive scoring (GAS) framework (Creal et al., 2013; Harvey and Chakravarty, 2008), where observation-driven dynamics is derived directly based on the likelihood function (5) of $\{Y_{ti}\}$, as is done in Massacci (2016) and Zhang and Schwaab (2017). However, one potential drawback of this approach is that the evolution of (σ_{ti}, ξ_{ti}) is then driven by the censored time series $\{Y_{ti}\}$ instead of the observed time series $\{X_{ti}\}$ and thus may suffer from potential information loss. Moreover, note that the scale parameter σ_{ti} of Y_{ti} is inherited from X_{ti}, thus a GAS-based dynamics for σ_{ti}

creates a discrepancy with the widely accepted GARCH literature where it is X_{ti} instead of Y_{ti} that drives the evolution of σ_{ti}.

Based on the above discussions, we seek alternatives for the dynamics of (σ_{ti}, ξ_{ti}). For the scale parameter σ_{ti}, since X_{ti} is directly observed, following the GARCH literature, we set

$$\sigma_{ti}^2 = \psi_0 + \psi_1 \sigma_{t-1,i}^2 + \psi_2 X_{t-1,i}^2,$$

which truthfully captures the dynamics of σ_{ti} for Y_{ti} and avoids the potential information loss caused by using censored time series $\{Y_{ti}\}$.

For dynamics of the tail index ξ_{ti}, following Zhao et al. (2018), we assume that

$$\log \xi_{ti} = \phi_0 + \phi_1 \log \xi_{t-1,i} + \phi_2 \exp(-|X_{t-1,i}|),$$

where the $\log$-transformation assures the positivity of ξ_{ti}. Note that ϕ_2 controls the relationship between ξ_{ti} and $|X_{t-1,i}|$. If $\phi_2 > 0$, a larger $|X_{t-1,i}|$ implies a smaller ξ_{ti} and a heavier tail for Y_{ti} (hence X_{ti}), creating the volatility/extreme-value clustering effects exhibited by financial time series. As is demonstrated in Zhao et al. (2018), the exponential function $\exp(-|x|)$ can be seen as a simplified version of the widely used logistic function $L/(1+a_0 \exp(-a_1 x))$ (Hansen, 1994; Boutahar et al., 2008; Hall et al., 2016) and offers flexible modeling for the tail index ξ_{ti}.

Together, the dynamic univariate Peak-over-Threshold (uPoT) model for $\{Y_{ti}\}$ is

$$F_{Y_{ti}}(y) = 1 - \left(\tau_i / \sigma_{ti}\right)^{-\xi_{ti}} \left(1 + y/\tau_i\right)^{-\xi_{ti}}, \quad (6)$$

$$\sigma_{ti}^2 = \psi_0 + \psi_1 \sigma_{t-1,i}^2 + \psi_2 X_{t-1,i}^2, \quad (7)$$

$$\log \xi_{ti} = \phi_0 + \phi_1 \log \xi_{t-1,i} + \phi_2 \exp(-|X_{t-1,i}|). \quad (8)$$

Importantly, by design, the observed time series $\{X_{ti}\}$ and censored time series $\{Y_{ti}\}$ share the same time-varying scale parameter σ_{ti} and tail index ξ_{ti}. Thus, the proposed dynamic uPoT model provides a semiparametric framework to study the marginal tail behavior of $\{X_{ti}\}$, without imposing strong parametric assumptions.

Compared to existing literature, there are several unique advantages of the newly proposed dynamic uPoT model. First of all, uPoT is sensible and interpretable as it is rooted on and directly derived from an underlying time series model that accurately reflects the characteristics of financial time series $\{X_{ti}\}$, such as volatility clustering. Second, in line with the extreme value literature, uPoT is scale-invariant in the sense that the tail index ξ_{ti} will not be affected by the unit of X_{ti} (hence Y_{ti}), thanks to the presence of the scale parameter σ_{ti} in (6). In contrast, Massacci (2016) and Zhang and Schwaab (2017) assume a dynamic univariate PoT model that takes the form $F_{Y_{ti}}(y) = 1 - (1+\tau_i)^{-\xi_{ti}} \left(1 + y/\sigma_{ti}\right)^{-\xi_{ti}}$. Due to the $(1+\tau_i)^{-\xi_{ti}}$ term, the estimated ξ_{ti} may change substantially if we perform the same analysis on, say $100X_{ti}$ (percentage scale) instead of X_{ti} (original scale), which is not ideal and may cause confusion for practitioners. Third, both dynamics of (σ_{ti}, ξ_{ti}) in uPoT are autoregressive and directly built upon the observed time series X_{ti}, which closely aligns with the GARCH literature and avoids information loss. As pointed out by Harvey (2013), if the volatility dynamics is not properly accounted for, the estimated time-varying behavior of the tail index ξ_{ti} can be potentially confounded with movements of the scale parameter σ_{ti} and thus renders erroneous conclusion. For comparison, Massacci (2016) proposes an autoregressive dynamics only for the tail index ξ_{ti} and instead models σ_{ti} as a function of $\xi_{t-1,i}$, which bears the risk of confounding dynamics of the tail index and volatility.

Remark 1. *An implicit restriction of the dynamic uPoT model is that the threshold τ_i needs to be larger than the scale parameter σ_{ti}, which is required by the validity*

of the cdf $F_{Y_{ti}}$ *in* (5)*. This is a mild restriction as the threshold* τ_i *is typically set to be an extreme quantile (e.g. 90%, 95%) of the time series* $\{X_{ti}\}$ *for an accurate tail approximation in* (4)*, making the scenario where* τ_i *is smaller than* σ_{ti} *unlikely. To further increase the flexibility of the dynamic uPoT model, we can keep the constant* k_i *in* (4) *and thus introduce an extra parameter* k_i *in* (6) *such that*

$$F_{Y_{ti}}(y) = 1 - k_i \left(\tau_i / \sigma_{ti}\right)^{-\xi_{ti}} \left(1 + y / \tau_i\right)^{-\xi_{ti}}.$$

We only require $k_i > 0$ *and thus relax the aforementioned restriction between* τ_i *and* σ_{ti}*. Note that this extension naturally nests the proposed dynamic uPoT as a special case since* $k_i \equiv 1$ *for the dynamic uPoT. In Section §5 of the supplementary material, we conduct a comparative study based on the extended model and illustrate that the empirical finding based on the extended model is highly consistent with the dynamic uPoT model. For model parsimony and numerical stability, in this paper we stick to dynamic uPoT with* $k_i \equiv 1$ *and leave the full examination of the extended model for future research.*

2.3 Dynamic bivariate PoT model

In this section, we propose the dynamic bivariate PoT (bPoT) model by further specifying the extreme value copula C_t in (2). We assume C_t to be a Gumbel copula, which is the most widely used extreme value copula, such that

$$C_t(u_1, u_2; \alpha_t) = \exp\left(-\left((-\log u_1)^{\alpha_t} + (-\log u_2)^{\alpha_t}\right)^{1/\alpha_t}\right),$$

where $\alpha_t \in [1, \infty)$ and regulates the dependence structure of C_t. When $\alpha_t = 1$, C_t becomes the independence copula and when $\alpha_t \to \infty$, C_t converges to the comonotonicity copula.

A popular measurement for quantifying the joint tail behavior of $\mathbf{X}_t$ (and thus $\mathbf{Y}_t$) is the tail dependence coefficient (Sibuya, 1959), where

$$\lambda_t = \lim_{q \nearrow 1} P\left(X_{t1} > \mathrm{VaR}_{X_{t1}}(q) \mid X_{t2} > \mathrm{VaR}_{X_{t2}}(q) \right), \qquad (9)$$

provided the limit exists. Here, $\mathrm{VaR}_{X_{ti}}(q)$ denotes the qth quantile of X_{ti}. In other words, the tail dependence coefficient λ_t measures the conditional probability of one asset suffering large losses given that the other asset experiencing large losses. Thus λ_t provides a natural metric for tail connectedness of two financial assets/markets. It can be shown that for Gumbel copula $C_t(u_1, u_2; \alpha_t)$, the tail dependence coefficient λ_t of $\mathbf{X}_t$ (and thus $\mathbf{Y}_t$) is $\lambda_t = 2 - 2^{1/\alpha_t}$. In the following, we hence refer to α_t as the tail coefficient or tail dependence parameter.

We note that there are a number of alternative extreme value copulas available, see Boldi and Davison (2007), Cooley et al. (2010), Ballani and Schlather (2011) and references therein. However, as is discussed in Ledford and Tawn (1996), Dupuis and Tawn (2001) and Poon et al. (2004), the exact specification of the extreme value copula is less crucial, and the Gumbel copula is flexible and robust for modeling bivariate extreme dependence.

Based on (2), on day t, given $(\sigma_{t1}, \xi_{t1}, \sigma_{t2}, \xi_{t2}, \alpha_t)$, the cdf of $\mathbf{Y}_t$ can be readily derived as

$$F_{\mathbf{Y}_t}(y_1, y_2) = C_t(F_Y(y_1; \tau_1, \sigma_{t1}, \xi_{t1}), F_Y(y_2; \tau_2, \sigma_{t2}, \xi_{t2}); \alpha_t),$$

where $F_Y(y; \tau, \sigma, \xi) = 1 - (\tau/\sigma)^{-\xi}(1 + y/\tau)^{-\xi}$ as defined in Section 2.2. The probability density function of $\mathbf{Y}_t$ is

$$f_{\mathbf{Y}_t}(y_1, y_2 \mid \sigma_{t1}, \xi_{t1}, \sigma_{t2}, \xi_{t2}, \alpha_t) = F_{\mathbf{Y}_t}(0,0) I(y_1 = 0, y_2 = 0) + \frac{\partial}{\partial y_1} F_{\mathbf{Y}_t}(y_1, 0) I(y_1 > 0, y_2 = 0)$$
$$+ \frac{\partial}{\partial y_2} F_{\mathbf{Y}_t}(0, y_2) I(y_1 = 0, y_2 > 0) + \frac{\partial^2}{\partial y_1 \partial y_2} F_{\mathbf{Y}_t}(y_1, y_2) I(y_1 > 0, y_2 > 0). \qquad (10)$$

The detailed formula of $f_{\mathbf{Y}_t}(y_1, y_2 \mid \sigma_{t1}, \xi_{t1}, \sigma_{t2}, \xi_{t2}, \alpha_t)$ is given in the supplementary material. In the following, for notational simplicity, we denote (10) as $f_{\mathbf{Y}_t}(y_1, y_2)$.

To capture the possibly time-varying behavior of joint tail risk, we further design dynamics for the tail coefficient parameter α_t. To ensure $\alpha_t \in [1, \infty)$, we set $\alpha_t = 1 + \exp(\gamma_t)$ and design an autoregressive process for $\gamma_t = \log(\alpha_t - 1)$ such that

$$\gamma_t = \beta_0 + \beta_1 \gamma_{t-1} + \beta_2 u_{t-1},$$

where $u_{t-1} \in \mathcal{F}_{t-1}$ is the observation-driven dynamics for γ_t. Note that we implicitly assume the parametric form of the Gumbel copula stays the same while the dependence parameter α_t is time-varying. We leave the possibility of parametric form changes for future research.

Unlike the time-varying marginal behavior (σ_{ti}, ξ_{ti}), the tail coefficient α_t is of higher order and there is no clear intuition or empirical evidence on how to design the observation-driven term u_t for its dynamics. To avoid subjective bias, we employ the GAS framework (Harvey and Chakravarty, 2008; Creal et al., 2013), under which the dynamic term u_t is derived based on the score of the likelihood function (10) of $\mathbf{Y}_t$, such that u_t takes the general form

$$u_t = S_t \cdot \frac{\partial \log f_{\mathbf{Y}_t}(y_{t1}, y_{t2})}{\partial \gamma_t} = S_t \cdot \frac{\partial \log f_{\mathbf{Y}_t}(y_{t1}, y_{t2})}{\partial \alpha_t} \cdot \frac{\partial \alpha_t}{\partial \gamma_t},$$

where S_t is a scaling function. In other words, u_t is a scaled-version of the first order derivative of the log-likelihood function of $\mathbf{Y}_t$. There are various ways to specify the scaling term S_t. For simplicity and numerical stability, following Creal et al. (2013) and Eckernkemper (2017), we set S_t to the identity function and thus $S_t = 1$. The detailed formula of u_t can be found in the supplementary material.

Together, the dynamic bivariate Peak-over-Threshold (bPoT) model for the censored bivariate time series $\{\mathbf{Y}_t\}$ is

$$F_{\mathbf{Y}_t}(y_1, y_2) = C_t(F_Y(y_1; \tau_1, \sigma_{t1}, \xi_{t1}), F_Y(y_2; \tau_2, \sigma_{t2}, \xi_{t2}); \alpha_t), \qquad (11)$$

$$\gamma_t = \beta_0 + \beta_1 \gamma_{t-1} + \beta_2 \frac{\partial \log f_{\mathbf{Y}_{t-1}}(Y_{t-1,1}, Y_{t-1,2})}{\partial \gamma_{t-1}}, \qquad (12)$$

where the marginal tail behavior of Y_{ti} (hence X_{ti}) is characterized by the dynamic uPoT model with time-varying scale parameter σ_{ti} and tail index ξ_{ti}, and the joint tail behavior of $\mathbf{Y}_t$ (hence $\mathbf{X}_t$) is captured by the dynamic tail coefficient α_t.

To our best knowledge, this is the first multivariate EVT-based time series model in the study of dynamic tail risk. It naturally generalizes the tail risk literature from univariate/marginal dimension to multivariate/joint dimension, and nests the univariate dynamic tail risk models in Massacci (2016) and Zhang and Schwaab (2017) as special cases. We emphasize the semiparametric nature of the dynamic bPoT model, as it is built upon assumptions only on the tail behavior of $\mathbf{X}_t$, instead of the entire distribution, and thus is robust against model misspecification.

Remark 2. *As is discussed by Ledford and Tawn (1996) and Poon et al. (2004), extreme value copula has the potential drawback of being too coarse for describing tails of multivariate distributions with asymptotic independence sufficiently accurately. Alternative models can be found in Ledford and Tawn (1997) and Bortot et al. (2002). We leave extensions of the dynamic bPoT model in this direction for future research.*

3 Estimation and Prediction

In this section, we discuss parameter estimation and model-based prediction for the proposed dynamic bPoT model and further conduct simulation studies to

demonstrate the performance of the maximum likelihood estimator (MLE) under finite sample.

3.1 Estimation

The parameters of the dynamic bPoT model are $\boldsymbol{\theta} = \bigcup_{i=1}^{2}(\psi_{0i}, \psi_{1i}, \psi_{2i}, \phi_{0i}, \phi_{1i}, \phi_{2i}) \bigcup (\beta_0, \beta_1, \beta_2)$, where $\boldsymbol{\theta}_i = (\psi_{0i}, \psi_{1i}, \psi_{2i}, \phi_{0i}, \phi_{1i}, \phi_{2i})$ is the parameter of the dynamic uPoT model for univariate time series $\{Y_{ti}\}_{t=1}^{T}$, i = 1, 2, and $\boldsymbol{\theta}_3 = (\beta_0, \beta_1, \beta_2)$ denotes the parameter for the time-varying tail dependence (extreme value copula) of the bivariate time series $\{\mathbf{Y}_t\}_{t=1}^{T}$.

For inference, we use the censored likelihood approach proposed by Ledford and Tawn (1996), which is also demonstrated by Huser et al. (2016) and Zhang and Schwaab (2017) to have superior performance. Specifically, given the bivariate time series $\{\mathbf{X}_t = (X_{t1}, X_{t2})\}_{t=1}^{T}$, we set the threshold $\boldsymbol{\tau} = (\tau_1, \tau_2)$ as extreme quantiles, e.g. 90%, of the marginal observations $\{X_{ti}\}_{t=1}^{T}, i = 1, 2$, and obtain the censored time series $\{\mathbf{Y}_t = (Y_{t1}, Y_{t2})\}_{t=1}^{T}$ via $Y_{ti} = \max(X_{ti} - \tau_i, 0), i = 1, 2.$ The estimation of $\boldsymbol{\theta}$ is then based on the (censored) likelihood of $\{\mathbf{Y}_t\}_{t=1}^{T}$.

Given $\{\mathbf{X}_t\}_{t=1}^{T}$ and initial values of $(\sigma_{11}, \xi_{11}, \sigma_{12}, \xi_{12}, \alpha_1)$ at t = 1, for any fixed $\boldsymbol{\theta}$, we can recover the latent time-varying parameters $\{\sigma_{t1}, \xi_{t1}, \sigma_{t2}, \xi_{t2}, \alpha_t\}_{t=1}^{T}$ based on the dynamics of bPoT specified in $(7), (8)$ and (12). The log-likelihood of $\{\mathbf{Y}_t\}_{t=1}^{T}$ can be calculated via

$$L(\boldsymbol{\theta}) = \sum_{t=1}^{T} \log f(\mathbf{Y}_t \mid \mathcal{F}_{t-1}) = \sum_{t=1}^{T} \log f(\mathbf{Y}_t \mid \sigma_{t1}, \xi_{t1}, \sigma_{t2}, \xi_{t2}, \alpha_t), \quad (13)$$

where $f(\mathbf{Y}_t \mid \sigma_{t1}, \xi_{t1}, \sigma_{t2}, \xi_{t2}, \alpha_t)$ is specified in (10). We define the full MLE as $\boldsymbol{\theta} = \arg\max L(\boldsymbol{\theta})$.

The optimization of (13) can be challenging, especially for large sample size T, as the dimension of $\boldsymbol{\theta}$ is large. To ease computational burden, we can start with a two-stage MLE, where we estimate the marginal parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ in the first stage, and estimate $\boldsymbol{\theta}_3$ in the second stage.

In the first stage, given univariate time series $\{X_{ti}\}_{t=1}^{T}$ and initial values of (σ_{1i}, ξ_{1i}), for any fixed $\boldsymbol{\theta}_i$, we can recover $\{\sigma_{ti}, \xi_{ti}\}_{t=1}^{T}$ based on the dynamics of uPoT specified in (7) and (8), and estimate $\boldsymbol{\theta}_i$ via

$$\boldsymbol{\theta}_i = \arg\max_{\boldsymbol{\theta}_i} L_i(\boldsymbol{\theta}_i) = \arg\max_{\boldsymbol{\theta}_i} \sum_{t=1}^{T} \log f(Y_{ti} \mid \sigma_{ti}, \xi_{ti}),$$

where $f(Y_{ti} \mid \sigma_{ti}, \xi_{ti})$ is specified in (5), for $i = 1, 2.$ In the second stage, keep the estimated $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ fixed, we estimate $\boldsymbol{\theta}_3$ via

$$\boldsymbol{\theta}_3 = \arg\max_{\boldsymbol{\theta}_3} L(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3),$$

where $L(\cdot)$ is the likelihood function defined in (13).

Under correct model specification, both the two-stage MLE and full MLE are consistent and asymptotically normal (Newey and McFadden, 1994), where the two-stage MLE is computationally more efficient. We can use the two-stage MLE $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3)$ as the starting point for the optimization of the full MLE $\boldsymbol{\theta}$. Creal et al. (2013) suggests to use the inverse Hessian matrix of the log-likelihood function (13) to estimate the asymptotic covariance matrix of the MLE $\boldsymbol{\theta}$.

Remark 3. *The initial value of* $(\xi_{11}, \sigma_{11}, \xi_{12}, \sigma_{12}, \alpha_1)$ *can be obtained via an MLE based on a static version of the dynamic bPoT model, where we assume that* $\{\mathbf{Y}_t\}_{t=1}^{T}$ *follows the bivariate distribution in* (10) *with constant parameters* $(\sigma_1, \xi_1, \sigma_2, \xi_2, \alpha)$.

3.2 Prediction

The proposed dynamic bPoT model is not a generative model for the original time series $\{\mathbf{X}_t\}$, as it only specifies the behavior of $\mathbf{X}_t$ when $\mathbf{X}_t$ is over the high threshold $\boldsymbol{\tau}$. Thus, given historical observations $\{\mathbf{X}_t\}_{t=1}^{T}$, dynamic bPoT can not be directly used to predict future behavior of $\{\mathbf{X}_t, t \geq T+1\}$. This can limit the potential applications of dynamic PoT in financial risk management.

To circumvent this limitation, we adapt the strategy in Poon et al. (2004) and extend dynamic bPoT to a generative model for time series $\{\mathbf{X}_t\}$ in an semiparametric way. For the marginal behavior of X_{ti}, given the threshold τ_i and (σ_{ti}, ξ_{ti}), we have

$$
\begin{aligned}
&F_{X_i}(x) = P(X_{ti} \leq x) = P(X_{ti} \leq \tau_i)P(X_{ti} \leq x \mid X_{ti} \leq \tau_i) + P(X_{ti} > \tau_i)P(X_{ti} \leq x \mid X_{ti} > \tau_i) \\
&= P(Y_{ti} = 0)P(X_{ti} \leq x \mid X_{ti} \leq \tau_i) + P(Y_{ti} > 0)P(Y_{ti} \leq x - \tau_i \mid Y_{ti} > 0) \\
&= (1 - (\tau_i / \sigma_{ti})^{-\xi_{ti}})F_{X_i}^{tr}(x) + \left(\tau_i / \sigma_{ti}\right)^{-\xi_{ti}} \left(1 - \left(x / \tau_i\right)^{-\xi_{ti}}\right)_+,
\end{aligned}
\tag{14}
$$

where $F_{X_i}^{tr}(x)$ denotes the truncated distribution of X_{ti} at τ_i, $(x)_+ = \max(0, x)$ and the last equality follows by plugging in the cdf of Y_{ti} in (6). We later specify $F_{X_i}^{tr}(x)$ nonparametrically. For the joint behavior of $\mathbf{X}_t$, following Poon et al. (2004), we directly employ the estimated Gumbel copula of the dynamic bPoT model with parameter α_t. See Remark 2 for more detailed discussion on this issue.

Thus, given the model parameter $\boldsymbol{\theta}$, threshold $\boldsymbol{\tau}$, truncated distributions $F_{X_1}^{tr}(x)$, $F_{X_2}^{tr}(x)$, and initial value $(\sigma_{11}, \xi_{11}, \sigma_{12}, \xi_{12}, \alpha_1)$, we can simulate $\{\mathbf{X}_t\}$ based on the above described generative model. The simulation procedure is summarized in Algorithm 1.

Algorithm 1 Simulation/Prediction of $\{\mathbf{X}_t\}$

1: Given $\boldsymbol{\theta}, \boldsymbol{\tau}, F_{X_1}^{tr}(x), F_{X_2}^{tr}(x)$, and initial value $(\sigma_{11}, \xi_{11}, \sigma_{12}, \xi_{12}, \alpha_1)$

2: For $t = 1, \ldots, T$

3: Step 1: Simulate (U_{t1}, U_{t2}) from a Gumbel copula with α_t

4: Step 2: Generate $X_{t1} = F_{X_1}^{-1}(U_{t1}; \sigma_{t1}, \xi_{t1})$ and $X_{t2} = F_{X_2}^{-1}(U_{t2}; \sigma_{t2}, \xi_{t2})$ based on (14)

5: Step 3: Update $(\sigma_{t+1,1}, \xi_{t+1,1}, \sigma_{t+1,2}, \xi_{t+1,2})$ based on (7) and (8)

6: Step 4: Set $Y_{t1} = \max(X_{t1} - \tau_1, 0), Y_{t2} = \max(X_{t2} - \tau_2, 0)$, update α_{t+1} based on (12)

7: Output $\{(X_{t1}, X_{t2})\}_{t=1}^{T}$

Given threshold $\boldsymbol{\tau}$ and historical observations $\{\mathbf{X}_t\}_{t=1}^{T}$, we specify the truncated distribution $F_{X_i}^{tr}(x)$ nonparametrically based on the empirical distribution of $\{X_{ti}\}_{t=1}^{T}$ such that

$$F_{X_i}^{tr}(x) = \frac{\sum_{t=1}^{T} I(X_{ti} \le \min(x, \tau_i))}{\sum_{t=1}^{T} I(X_{ti} \le \tau_i)}. \qquad (15)$$

To make predictions for $\{\mathbf{X}_t, t \ge T+1\}$ in the future, we set the initial value at $(\hat{\sigma}_{T1}, \hat{\xi}_{T1}, \hat{\sigma}_{T2}, \hat{\xi}_{T2}, \hat{\alpha}_T)$, which can be recovered based on $\boldsymbol{\theta}$, and run Algorithm 1.

This provides a parametric bootstrap procedure for $\{\mathbf{X}_t, t \ge T+1\}$ conditional on historical observations $\{\mathbf{X}_t\}_{t=1}^{T}$. Based on the simulated samples, common tail risk measures used in financial risk management such as conditional Value-at-Risk, Expected Shortfall, can then be estimated for $\{\mathbf{X}_t, t \ge T+1\}$. For more details on risk measures, we refer to McNeil et al. (2005). In Section 5, we demonstrate the performance of the proposed prediction algorithm for tracking dynamics of joint tail risk in an empirical study.

Remark 4. *The proposed generative model of* $\{\mathbf{X}_t\}$ *needs to be used with caution. Specifically, the Gumbel copula is not suitable for modeling the joint*

behavior of $\mathbf{X}_t$ *over entire* $\mathbb{R}^2$*, but is only appropriate for modeling the joint behavior in the tail region. Thus, the prediction produced by Algorithm 1 should only be used for estimating/forecasting risk measures that track tail behavior of* $\mathbf{X}_t$ *such as Value-at-Risk or Expected Shortfall at high quantiles, but not for risk measures such as variance. See Poon et al. (2004) for more detailed discussion.*

3.3 Simulation study

In this section, we conduct numerical experiments to demonstrate the performance of MLE under finite sample. We set the true parameter of the dynamic bPoT model as

$$\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2 = (0.033, 0.816, 0.005, -0.017, 0.949, 0.112) \text{ and } \boldsymbol{\theta}_3 = (0.01, 0.99, 0.1),$$

which closely resembles the estimated parameters from the real data study reported in Table 3 and Table 5 in Section 4, where dynamic bPoT is employed to study the dynamics of joint tail behavior of global financial markets.

To further reflect the characteristics of financial time series, we define $F^{tr}_{X_i}(x)$ nonparametrically based on the empirical distribution of the 6202 daily negative log-returns of Dow Jones Industrial Average Index from 1990 to 2018, which is part of the data used in the empirical study in Section 4. Denote the data of daily negative log-returns as $\{r_t\}_{t=1}^{6202}$. We set $\tau_1 = \tau_2 = 1.12$, which is the 90% sample quantile of $\{r_t\}_{t=1}^{6202}$, and we set $F^{tr}_{X_i}(x) = \sum_{t=1}^{6202} I(r_t \le \min(x, \tau_i)) / \sum_{t=1}^{6202} I(r_t \le \tau_i)$ for $i = 1, 2.$

Given the true parameter $\boldsymbol{\theta}$, threshold $\boldsymbol{\tau}$ and the truncated distributions $F^{tr}_{X_1}(x), F^{tr}_{X_2}(x)$, we simulate the bivariate time series $\{\mathbf{X}_t\}_{t=1}^{T}$ using Algorithm 1. We set T = 3000, 6000, 9000 and conduct 500 simulations for each sample size T. Given the simulated bivariate time series $\{\mathbf{X}_t\}_{t=1}^{T}$, we set $\tilde{\tau}_i =$ 90% quantile of $\{X_{ti}\}_{t=1}^{T}$ and obtain $\{Y_{ti} = \max(X_{ti} - \tilde{\tau}_i, 0)\}_{t=1}^{T}$ for i = 1, 2. We estimate $\boldsymbol{\theta}$ using both full MLE and two-stage MLE as defined in Section 3.1.

The estimation result is summarized in Table 1. Since $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$, we only report the result for $\boldsymbol{\theta}_1$ due to limited space. It is clear that both full MLE and two-stage MLE provide satisfactory estimation accuracy and their performance improves as T grows. The performance of two-stage MLE is slightly worse than full MLE within a comparable range.

For each experiment, based on the estimated $\boldsymbol{\theta}$, we can recover the latent time-varying parameters $\{(\hat{\sigma}_{t1}, \hat{\xi}_{t1}, \hat{\sigma}_{t2}, \hat{\xi}_{t2}, \hat{\alpha}_t)\}_{t=1}^{T}$. To further evaluate the accuracy of the recovered dynamics, we report in Table 2 the average Pearson correlation and Spearman's rho[1] between the true and estimated time-varying parameters across the 500 experiments. We skip the result for $(\hat{\sigma}_{t2}, \hat{\xi}_{t2})$ as $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2$. As can be seen, the recovered dynamics based on both full MLE and two-stage MLE are highly correlated with the true dynamics, with both the Pearson correlation and Spearman's rho approaching 1 quickly as T increases. Again, the result based on full MLE is better than two-stage MLE within a comparable range.

4 Dynamics of Tail Connectedness for Global Stock Markets

In this section, we study the dynamics of joint tail risk across global stock markets. We use the daily negative log-returns from nine major stock indices across three continents: North America with Dow Jones Industrial Average Index (DJI, U.S.), S&P 500 Index (SP500, U.S.) and Nasdaq Composite Index (NASDAQ, U.S.), Europe with FTSE 100 Index (FTSE, U.K.), CAC 40 Index (CAC, France), DAX Index (DAX, Germany) and Euro Stoxx 50 Index (ES50, Euro zone), and Asia with Nikkei 225 Index (N225, Japan) and Hang Seng Index (HSI, Hong Kong).

The sample period is from March 1990 to February 2018, with a total of T = 6202 daily observations. Similar to Massacci (2016), we account for holidays by deleting observations from trading days for which at least one market has a return identically equal to zero. Following Poon et al. (2004), we address the

synchronization issue of international equity markets by moving observations of the three U.S. indices one day back in the data.

We set the threshold τ_i to be the 90% quantile of negative log-returns of each index $\{X_{ti}\}_{t=1}^{6202}, i = 1,\ldots,9$, which is shown to work well in Chavez-Demoulin et al. (2005), Chavez-Demoulin et al. (2014), and Massacci (2016), and set $Y_{ti} = \max(X_{ti} - \tau_i, 0)$. In Section §4.1 of the supplementary material, we further perform the same analysis with the threshold τ_i being set at 85% or 95% quantile. We find that the analysis result is highly consistent across different thresholds τ_i and thus confirms the robustness of the empirical finding by the proposed model.

4.1 Results for dynamics of marginal tail risk

We first model the nine stock indices individually using the proposed dynamic uPoT model. Table 3 summarizes the estimation result for each stock index. As can be seen, for all nine indices, the estimated autoregressive parameter ψ_{i1} and ϕ_{i1} for the time-varying scale parameter σ_{ti} and tail index ξ_{ti} both are close to unity and are highly significant, revealing that marginal tail risk is persistent in stock markets. This finding agrees with Massacci (2016), where the author studies the dynamics of marginal tail risk of the G7 countries MSCI indices.

Given the estimated dynamic uPoT model, we can recover the latent time-varying parameters $\{(\hat{\sigma}_{ti}, \hat{\xi}_{ti})\}$ for each stock index $i = 1,\ldots,9.$ For illustration purposes, Figure 1(b) plots the time series of the average tail index $\{\bar{\hat{\xi}}_t = \sum_{i=1}^{9} \hat{\xi}_{ti} / 9\}$ among nine indices. As can be seen, the tail index varies substantially across time and reaches its lowest level around 2008-2009, during the financial crisis. This is in line with the findings in Kelly and Jiang (2014), Massacci (2016), Mao and Zhang (2018), and Zhao et al. (2018), that the marginal tail behavior of financial markets is time-varying and the market exhibits heavier tails during crisis.

To further demonstrate the dynamic uPoT's ability to capture the time-varying volatility of the observed time series $\{X_{ti}\}$, we additionally fit a GARCH(1,1) model for each time series $\{X_{ti}\}$ and recover the estimated volatility $\{\tilde{\sigma}_{ti}\}$ by GARCH. For each stock index, we calculate the (Pearson) correlation between $\{\hat{\sigma}_{ti}\}_{t=1}^{6202}$ and $\{\tilde{\sigma}_{ti}\}_{t=1}^{6202}$, and report the result in Table 3. To rule out spurious correlation, we additionally report the (Pearson) correlation between the first-order differenced series $\{\Delta\hat{\sigma}_{ti} = \hat{\sigma}_{ti} - \hat{\sigma}_{t-1,i}\}_{t=2}^{6202}$ and $\{\Delta\tilde{\sigma}_{ti} = \tilde{\sigma}_{ti} - \tilde{\sigma}_{t-1,i}\}_{t=2}^{6202}$ in Table 3. The correlation ranges from 0.77 to 0.98 with an average correlation of 0.932 across the nine indices, indicating that the dynamic scale parameter $\{\hat{\sigma}_{ti}\}$ is an accurate measure of the underlying volatility. This is further confirmed by Figure 1(a), where we plot (in standardized scale for comparison) the average scale parameter $\{\hat{\bar{\sigma}}_t = \sum_{i=1}^{9} \hat{\sigma}_{ti} / 9\}$ estimated by dynamic uPoT and average volatility $\{\tilde{\bar{\sigma}}_t = \sum_{i=1}^{9} \tilde{\sigma}_{ti} / 9\}$ estimated by GARCH(1,1) across the nine indices. The two series are seen to highly agree with each other. As discussed in Section 2.2, compared to existing literature, this is a unique advantage of the dynamic uPoT model, which helps prevent the potential confounding between dynamics of volatility and tail index.

4.2 Results for dynamics of joint tail risk

We now present the estimation result for dynamics of joint tail behavior between global stock indices. There is a total of 36 bivariate pairs from the nine stock indices and we model each pair of indices with the dynamic bPoT model. To investigate whether the joint tail behavior of each pair exhibits time-varying dynamics, we perform a formal hypothesis test, where for each pair we test

$$H_0 : \beta_1 = \beta_2 = 0.$$

Under H_0, the pair of indices has a constant tail coefficient $\alpha_t \equiv \beta_0$. Note that H_0 is nested within the dynamic bPoT model, thus can be examined via the classical likelihood ratio test.

Each test is conducted at Type-I error $= 0.05/36$ level so that we control the family-wise error rate to be under 0.05. Among the $k = 1,\ldots 36$ pairs, 13 pairs have statistically significant time-varying joint tail behavior. In Table 4, we report the 13 pairs along with its average tail coefficient $\hat{\bar{\alpha}}_k = \sum_{t=1}^{6202} \hat{\alpha}_{tk} / 6202$, median tail coefficient $\hat{\underline{\alpha}}_k = \text{median}\{\hat{\alpha}_{tk}\}_{t=1}^{6202}$ and χ^2 -value for the likelihood ratio test. We note that the average tail coefficient $\hat{\bar{\alpha}}_k$ and median tail coefficient $\hat{\underline{\alpha}}_k$ are highly consistent with each other. Notably, for all pairs of stock indices within the same continent, the testing result supports the presence of time-varying joint tail risk. Interestingly, the three inter-continental pairs with statistically significant time-varying tail coefficient are between N225 and the three U.S. indices, indicating a dynamic nature of joint tail behavior between the Japanese and U.S. stock market.

For pairs with statistically insignificant joint tail risk, a static tail coefficient $\hat{\alpha}_k$ is estimated instead. Due to limited space, for both significant and insignificant pairs, we rank the pairs by their (average) tail coefficient and only report the estimation results for the top three and bottom three pairs. The reported result is representative and the full result is available upon request. Table 5 summarizes the selected estimation result for significant pairs. Similar to the dynamics of marginal tail behavior, the joint tail risk is also highly persistent as the estimated autoregressive parameter β_1 for the time-varying tail coefficient α_t is close to unity.

The selected estimation result for insignificant pairs is summarized in Table 6, with all pairs being inter-continental. Compared to the significant pairs, the magnitude of tail coefficients is much smaller for the insignificant pairs, with the

largest being 1.259, indicating much weaker tail dependence. The complete result of the estimated tail coefficients for insignificant pairs is reported in Table S.2 of the supplementary material. Interestingly, the results in Table S.2 indicate that among inter-continental pairs, stock indices across the Atlantic have the lowest level of tail dependence. In other words, European and U.S. financial markets have the smallest joint tail risk, and may work best for tail risk diversification.

To summarize, the general pattern of joint tail risk for global stock markets is that for stock indices from the same continent, the joint tail risk is persistent, time-varying and operates at a relatively high level, while for markets across different continents, the joint tail risk stays roughly constant and operates at a relatively low level. This is intuitive and agrees with the findings in Poon et al. (2004) and Hoga (2018).

4.3 New measure for tail connectedness

Given a set of risky assets, a natural question to ask is how *connected* they are. Several measures of connectedness are proposed to answer the question. Kritzman et al. (2011) and Billio et al. (2012) build connectedness measures for financial returns via principle component analysis (PCA). Diebold and Yilmaz (2014) build several measures from variance decomposition to study realized volatilities. Built upon the idea of Billio et al. (2012), Massacci (2016) proposes a tail connectedness measure by applying PCA to the marginal tail indices of individual assets. We contribute to the literature by defining a new tail connectedness measure via the dynamic bPoT model, which is more natural, intuitive and interpretable.

Based on the estimated dynamic bPoT, we can recover the time-varying tail coefficient parameter $\hat{\alpha}_{tk}$ for each pair of stock indices, which in turn recovers the time-varying tail dependence between two stock markets via $\hat{\lambda}_{tk} = 2 - 2^{1/\hat{\alpha}_{tk}}$. As is

defined in (9) in Section 2.3, the tail dependence coefficient $\hat{\lambda}_{tk}$ measures the conditional probability of one market suffering large losses given that the other market experiencing large losses. Thus, by nature, $\hat{\lambda}_{tk}$ serves as an interpretable and intuitive measure of tail connectedness between two financial markets. Treating $\hat{\lambda}_{tk}$ as a similarity measure, we recover a time-varying network of the nine stock markets. In the following, we analyze the behavior of the network to gain more insights on the tail connectedness of global stock markets.

We first examine the strength of tail connectedness between different stock indices. For each pair of indices $k = 1, \ldots, 36$, we calculate its average tail dependence coefficient $\bar{\hat{\lambda}}_k = \sum_{t=1}^{6202} \hat{\lambda}_{tk} / 6202$ in the observation period. Figure 2(a) plots the histogram of $\{\bar{\hat{\lambda}}_k\}_{k=1}^{36}$. The histogram can be clearly divided into two clusters with one cluster having relatively high tail connectedness (~0.65) and one cluster having low tail connectedness (~0.2). Interestingly, all the pairs in the high tail connectedness cluster come from the same continent and all the pairs in the low tail connectedness cluster are inter-continental, with (HSI, N225) an exception (the rightmost bin of the left cluster). Figure 2(b) provides the boxplot of $\{\bar{\hat{\lambda}}_k\}_{k=1}^{36}$ grouped by within-continental pairs and inter-continental pairs, which further implies that the joint tail risk is much lower for inter-continental stock markets compared to within-continental ones. This again indicates that to hedge against joint tail risk, the diversification can be maximized by allocating resources across continents. For robustness check, in Section §4.2 of the supplementary material, we further present the analysis result based on the median tail dependence coefficient $\underline{\hat{\lambda}}_k = \text{median}\{\hat{\lambda}_{tk}\}_{t=1}^{6202}$ in the observation period. The result is highly consistent with Figure 2 and thus confirms the difference between the joint tail risk of inter-continental and within-continental stock markets.

We now study the dynamics of tail connectedness among the global stock market. For each day $t = 1, \ldots, 6202$, we calculate the average tail dependence $\hat{\bar{\lambda}}_t = \sum_{k=1}^{36} \hat{\lambda}_{tk} / 36$ across all pairs of stock indices, which is defined to be the tail connectedness measure of the global stock market on day t (given by dynamic bPoT). For comparison, we also assess the tail connectedness using two alternative methods. The first method uses the same tail connectedness measure as dynamic bPoT but is based on a nonparametric moving window estimation. Specifically, for each day t, a static bPoT is estimated for each pair of stock indices based on observations $\{Y_{si}, s = t - h_1 + 1, \ldots, t\}_{i=1}^{9}$, where h_1 is the window size. For numerical stability, we set $h_1 = 1000$, which corresponds to roughly 4 years of observations. Based on the moving window estimation on each day, we can similarly calculate the average tail dependence $\hat{\bar{\lambda}}_t$. The second method is the PCA-based measure proposed in Massacci (2016). Denote $\hat{\Sigma}_t$ as the sample covariance matrix of the marginal tail indices $\{\hat{\xi}_{ti}\}_{i=1}^{9}$, for each day t, the tail connectedness measure is defined to be the ratio between the maximum eigenvalue of $\hat{\Sigma}_t$ and the sum of all eigenvalues of $\hat{\Sigma}_t$. To estimate $\hat{\Sigma}_t$, we again need a period of observations $\{\hat{\xi}_{si}, s = t - h_2 + 1, \ldots, t\}_{i=1}^{9}$, where h_2 is the window size. Following Billio et al. (2012), we set $h_2 = 756$, which corresponds to 36 months of observations. For comparison purposes, we also set $h_2 = 100$, which is used in Massacci (2016).

Figure 3(a) plots the tail connectedness measures given by dynamic bPoT, the moving window method and the PCA ($h_2 = 756$) method. For better comparison, each series is standardized. In general, the three series share similar behavior, with the Pearson correlation (Spearman's rho) between bPoT and PCA being 0.79 (0.81), between PCA and moving window being 0.81 (0.76), and between bPoT and moving window being 0.77 (0.72). All three measures reach their peak around the financial crisis period, indicating that diversification in global stock

markets may not hedge against joint tail risk during the crisis period, which resembles the findings in Diebold and Yilmaz (2014) and Massacci (2016). Figure 3(b) plots the tail connectedness measure given by PCA ($h_2 = 100$), which seems to be quite different from PCA ($h_2 = 756$), suggesting that the PCA measure may be sensitive to the choice of window size.

The analysis demonstrates that the newly proposed tail connectedness measure based on dynamic bPoT can accurately capture the time-varying joint tail behavior of global financial markets. Compared to the univariate approach in Massacci (2016), dynamic bPoT directly measures the time-varying tail connectedness via joint modeling, which is more intuitive and interpretable. Moreover, the new measure does not depend on a tuning parameter such as the window size and thus can track real-time dynamics without any lag. Also, unlike the univariate approach which only provides an overall measure of tail connectedness across all markets, dynamic bPoT gives a measure of tail-connectedness for each pair of stock indices and thus helps assess joint tail risk dynamics at a finer scale.

5 Portfolio Optimization for S&P500 Sector Indices

In this section, we demonstrate the applications of the dynamic bPoT model in portfolio optimization, expanding the application of tail risk literature to the area of active financial risk management. Through an empirical example, we show that the dynamic bPoT model can accurately predict joint tail risk and help construct more risk-rewarding portfolios.

The data we consider here is the daily negative log-returns of nine sector indices of the S&P 500 index[2], where each sector index tracks the performance of S&P500 in a specific sector according to the Global Industry Classification Standard. The nine sectors are Industrial (LI), Energy (LE), Materials (LB), Consumer Discretionary (LY), Financial (LF), Technology (LK), Health Care (LV), Consumer Staples (LP), and Utilities (LU). The observation period is from

January 1999 to December 2017, which consists of T = 4780 days and roughly 956 trading weeks.

To investigate the joint tail behavior of the nine sector indices, we first fit the dynamic bPoT model for each pair of sector indices. Notably, *all* 36 pairs of sector indices exhibit statistically significant time-varying tail coefficient, which is in line with the finding in Section 4 that joint tail behavior between within-continental indices is time-varying. See Section §3.2 of the supplementary material for more detailed estimation results.

5.1 Dynamic bPoT-based portfolio optimization

We now propose a dynamic bPoT-based optimal portfolio construction framework. Roughly speaking, in the proposed framework, we manage a portfolio that consists of the nine sector indices. At the end of each trading week[3], say day *t*, we seek to forecast an optimal portfolio weight $\mathbf{w}_t = \{\hat{w}_{ti}\}_{i=1}^{9}$ for the next trading week of days $(t+1, t+2, \ldots, t+5)$ such that a certain risk measure (e.g. variance, Value-at-Risk, Expected Shortfall) of the portfolio with weight $\mathbf{w}_t$ is minimal among all possible portfolios. Mathematically speaking, denote the portfolio weight as $\mathbf{w} = \{w_i\}_{i=1}^{9}$, we aim to solve the following constrained optimization problem:

$$\min_{\mathbf{w}} \Phi_t(\mathbf{w}), \text{ where } \mathbf{w} \geq 0, \mathbf{w}'\mathbf{1} = 1. \qquad (16)$$

where $\Phi_t(\mathbf{w})$ denotes the selected risk measure of the portfolio with weight $\mathbf{w}$ in the next trading week of days $(t+1, t+2, \ldots, t+5)$. The minimizer $\mathbf{w}_t$ is the optimal portfolio weight.

To operationalize (16), we need an estimation of the objective function $\Phi_t(\mathbf{w})$ for a given $\mathbf{w}$. The typical approach in the literature is to estimate $\Phi_t(\mathbf{w})$ by a statistical model that jointly models all portfolio components, see for example

Harris and Mazibas (2013) and Zhao and Zhang (2018). However, in our setting, this approach is infeasible as we do not have a joint multivariate model for all nine sector indices but only bivariate models for each pair. This motivates us to propose an alternative portfolio optimization strategy based on bivariate joint tail risk predicted by the dynamic bPoT models for each pair.

For each pair $k = 1, \ldots, 36$, denote $(p_1(k), p_2(k))$ as the bivariate sector indices (among $1, \ldots, 9$) that it consists of. Denote $\Phi_t(\mathbf{w}^k)$ as the risk measure of the portfolio with weight $\mathbf{w}^k$, where $\mathbf{w}^k = (w_1^k, \ldots, w_9^k)$ and $w_i^k = 0$ if $i \neq p_1(k)$ or $p_2(k)$. In other words, a portfolio with weight $\mathbf{w}^k$ *only* consists of the *k*th pair of sector indices.

For the *k*th pair, conditioned on its estimated dynamic bPoT model on day *t*, $\Phi_t(\mathbf{w}^k)$ can be approximated by parametric bootstrap. Specifically, using Algorithm 1 in Section 3.2, we generate B = 10000 bootstrap samples $\left\{\left(X^b_{t+j,p_1(k)}, X^b_{t+j,p_2(k)}\right), j = 1, \ldots, 5\right\}_{b=1}^{B}$ for the log-returns of the two indices $(p_1(k), p_2(k))$ on days $(t+1, \ldots, t+5)$. Consequently, $\Phi_t(\mathbf{w}^k)$ can be approximated by its empirical counterpart based on the B bootstrap portfolio returns $\left\{\sum_{j=1}^{5}\sum_{i=1}^{2} w_{p_i(k)} X^b_{t+j,p_i(k)}\right\}_{b=1}^{B}$. We then define the optimal weight $\mathbf{w}_t^k$ for the *k*th pair as

$$\mathbf{w}_t^k = \arg\min_{\mathbf{w}^k} \Phi_t(\mathbf{w}^k), \text{ where } \mathbf{w}^k \geq 0, (\mathbf{w}^k)'\mathbf{1} = 1, \qquad (17)$$

for $k = 1, \ldots, 36$.

Thus, on day *t*, for each pair of sector indices, we have its estimated tail coefficient $\hat{\alpha}_{tk}$, its optimal weight $\mathbf{w}_t^k$, and its minimized risk measures $\Phi_t(\mathbf{w}_t^k)$. We construct the optimal portfolio as follows. Denote $\mathcal{A}_t$ as the set of pairs with the 10 smallest[4] tail coefficients $\hat{\alpha}_{tk}$ and denote $\mathcal{B}_t$ as the set of pairs with the 10

smallest[4] minimized risk measure $\Phi_t(\mathbf{w}_t^k)$. We take $\mathcal{C}_t = \mathcal{A}_t \cap \mathcal{B}_t$ and define the optimal portfolio weight as

$$\mathbf{w}_t = \sum_{k \in \mathcal{C}_t} \mathbf{w}_t^k / |\mathcal{C}_t|,$$

where $|\mathcal{C}_t|$ is the cardinality of set $\mathcal{C}_t$. The heuristic of the proposed bivariate optimization strategy is that to minimize the tail risk of the portfolio, we select pairs of sector indices with weak tail dependence $\hat{\alpha}_{tk}$ and with less joint risk $\Phi_t(\mathbf{w}_t^k)$.

Remark 5. *To obtain the optimal weight $\mathbf{w}_t$, we need to solve k = 36 optimizations. However, note that the optimization for each $\Phi_t(\mathbf{w}^k)$ is a one-dimensional problem and thus can be efficiently solved. In other words, through the proposed bivariate optimization strategy, we decompose a 9-dimensional optimization problem into 36 one-dimensional problems that can be solved parallelly.*

5.2 Backtesting result

We employ a backtesting procedure to examine the performance of the weekly[5] optimal portfolio constructed by the bivariate optimization strategy. The basic steps are as follows. We initially estimate the dynamic bPoT models using the first *s* = 2000 observations. Based on the estimated models, we forecast the one-week ahead out-of-sample optimal portfolio weight $\mathbf{w}_s = \{\hat{w}_{si}\}_{i=1}^9$ for the next trading week of days $(s+1, s+2, \ldots, s+5)$ using the bivariate portfolio construction framework described above. The performance of the constructed optimal portfolio is $\hat{r}_s = \sum_{t=s+1}^{s+5} \sum_{i=1}^{9} \hat{w}_{si} r_{ti}$, where r_{ti} is the actual log-return of the *i*th sector index on day *t*. The estimation window is then rolled forward one trading week (5 days) and the optimal portfolio weight for the next week is generated. The last iteration

uses observations on days $\{T-2004,\ldots,T-5\}$ to generate optimal portfolio weight for the trading week of days $(T-4,T-3,\ldots,T)$.

We set the risk measure $\Phi_t(\mathbf{w}^k)$ to be the 99% Expected Shortfall (ES) of the cumulative portfolio returns in the next trading week of days $(t+1,\ldots,t+5)$. The reason why we set the level of ES close to 1 is that we only want to avoid the extreme loss of the portfolio instead of all the risk. The intuition is that if we try to eliminate all the risk, the constructed portfolio will also lack the ability to capture upward movements of the stock market. For more information on ES, readers are referred to McNeil et al. (2005).

The starting date of the out-of-sample test is January 03, 2007 and the ending date is December 29, 2017, consisting of 554 trading weeks. The test set covers the financial crisis period from 2007 to 2009 and the post-crisis market rally from 2010 to 2017, which enables us to examine the performance of the proposed portfolio optimization framework under different market scenarios. For comparison, we also report the performance of an equal-weight portfolio where $\mathbf{w}_t \equiv (1/9,\ldots,1/9)$ without any optimization, and the performance of the S&P500 index during the same period.

The summary statistics of the portfolio returns in the out-of-sample test set can be found in Table 7. Overall, compared to the equal-weight portfolio and the S&P500 index, the dynamic bPoT-based portfolio delivers higher annualized returns and considerably lower risk, in terms of standard deviation (SD), Value-at-Risk (VaR), and ES of the portfolio. Moreover, the portfolio constructed by dynamic bPoT offers better Sharpe ratio, which is the most commonly used measure for risk-adjusted returns in the financial industry[6]. Note that the performance of the equal-weight portfolio and the S&P500 index are very similar, indicating that the equal-weight portfolio can serve as an accurate market proxy.

To visualize the comparison, we plot the cumulative values of the equal-weight and bPoT-weight portfolios throughout the test set in Figure 4(a), where the portfolio value constructed by dynamic bPoT is seen to almost always stay on top. The vertical dotted line marks the date (Feb 26, 2009) when both portfolios reach its minimum level, and we call it the minimal point. The market scenario before the minimal point can be seen as in "crisis" state and the one after the minimal point can be seen as in post-crisis "normal/rally" state. To better assess portfolio performance under different market scenarios, separate analyses are performed on portfolio returns before and after the minimal point. We summarize the results in Table 7, where consistent with the overall result, the bPoT-based portfolio delivers similar or better returns with considerably smaller risk.

To further examine the dynamic bPoT model's ability of joint tail risk prediction, we conduct an additional backtesting for conditional VaR prediction. Specifically, for each pair $k = 1, \ldots, 36$, we consider the equal-weight portfolio (i.e. $w_{p_1(k)} \equiv w_{p_2(k)} \equiv 0.5$). On day t, we forecast the $q100\%$ VaR for the cumulative loss of the equal-weight portfolio over the next trading week of days $(t+1, t+2, \ldots, t+5)$ by the q-th quantile of the bootstrap cumulative loss $\left\{ -\sum_{j=1}^{5} \sum_{i=1}^{2} 0.5 X^{b}_{t+j, p_i(k)} \right\}_{b=1}^{B}$. Note that this is a challenging task as we aim to predict 5-day ahead VaR for a cumulative loss.

A violation happens when the actual cumulative weekly loss $r_t = -\sum_{s=t+1}^{t+5} \sum_{i=1}^{2} 0.5 r_{s, p_i(k)}$ is over the q-th quantile of the bootstrap sample, where q varies among 0.95, 0.99, 0.995 and 0.999. For each pair of sector indices $k = 1, \ldots, 36$, we record its realized VaR level over the 554 trading weeks in the test set. For $q = 0.95, 0.99, 0.995, 0.999$, the average realized VaR levels across the 36 pairs are 0.954, 0.991, 0.996 and 0.999, respectively. The boxplots of the realized VaR levels for the 36 pairs are given in Figure 4(b). As can be seen, the VaR predicted by the dynamic bPoT model performs reasonably well, especially for

high quantiles $q = 0.995$ and 0.999, with slightly over-prediction for $q = 0.95$ and 0.99. Together with the favorable performance in portfolio optimization, this further demonstrates the dynamic PoT model's ability to accurately track dynamics of joint tail risk and its promising potential in active financial risk management.

6 Conclusion

This paper proposes a novel dynamic bivariate PoT model to study the time-varying joint tail behavior of multivariate financial time series. The proposed model generalizes the existing literature on tail risk from univariate/marginal dimension to multivariate/joint dimension. We define a sensible and interpretable tail connectedness measure and examine the dynamics of joint tail behavior of global stock markets: empirical evidence suggests markets from the same continent have time-varying and high-level joint tail risk and tail connectedness increases during periods of crisis. We further propose a novel bivariate portfolio optimization procedure which expands the applications of tail risk literature to the area of active financial risk management. Backtesting results demonstrate the promising risk-rewarding performance of the proposed optimization procedure and the dynamic bPoT model's ability to accurately track time-varying joint tail risk. One future research direction is to extend the dynamic bivariate PoT model to higher dimension with the construction of a flexible multivariate extreme value copula.

Notes

[1]For definition of Spearman's rho, see Joe (2014). In Table S.1 of the supplementary material, we further report the average Pearson correlation between the first-order differenced true and estimated time-varying parameters across the 500 experiments, as a robustness check to rule out spurious correlation.

[2]In total, there are eleven sector indices of the S&P 500 index. The Real Estate and the Communication Services sector have relatively short history and thus are excluded from the analysis.

[3]Note that for demonstration, this is a weekly portfolio optimization procedure. Other optimization frequency, such as daily or biweekly, can be easily adapted.

[4]In the backtesting result, we find that the performance of the optimization procedure is robust to the cardinality of set $\mathcal{A}_t$ and $\mathcal{B}_t$. Here, we present the result where the cardinality is 10.

[5]As a robustness check of the effectiveness of the bPoT-based optimization and to further lower the potential transaction costs due to portfolio optimization, in Section §3.3 of the supplementary material, we further present the backtesting result for a biweekly optimized portfolio based on bPoT.

[6]For a given portfolio, its Sharpe ratio is defined to be $\frac{r - r_f}{\sigma}$, where r is its expected return, σ is its standard deviation, and r_f is the risk free rate. Here, we set r_f to be the federal funds rate, which is 0.00-1.75% during the test period. For simplicity we set $r_f = 0$.

References

Ballani, F. and Schlather, M. (2011). A construction principle for multivariate extreme value distributions. *Biometrika*, 98:633–645.

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104:535–559.

Boldi, M. O. and Davison, A. C. (2007). A mixture model for multivariate extremes. *Journal of the Royal Statistical Society - Series B*, 69(2):217–229.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Bortot, P., Coles, S. G., and Tawn, J. A. (2002). The multivariate gaussian tail model: An application to oceanographic data. *Journal of the Royal Statistical Society - Series C*, 49(1):31–49.

Bosma, J. J., Koetter, M., and Wedow, M. (2019). Too connected to fail? inferring network ties from price co-movements. *Journal of Business & Economic Statistics*, 37(1):67–80.

Boutahar, M., Dufrénot, G., and Péguin-Feissolle, A. (2008). A simple fractionally integrated model with a time-varying long memory parameter d_t. *Computational Economics*, 31:225–241.

Chavez-Demoulin, V., Davison, A. C., and McNeil, A. J. (2005). Estimating value-at-risk: a point process approach. *Quantitative Finance*, 5(2):227–234.

Chavez-Demoulin, V., Embrechts, P., and Sardy, S. (2014). Extreme-quantile tracking for financial time series. *Journal of Econometrics*, 188(1):44–52.

Cooley, D., Davis, R. A., and Naveau, P. (2010). The pairwise beta distribution: A flexible parametric multivariate model for extremes. *Journal of Multivariate Analysis*, 101(9):2103–2117.

Creal, D., Koopman, S., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795.

Creal, D., Koopman, S. J., and Lucas, A. (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics*, 29(4):552–563.

Diebold, F., Schuermann, T., and Stroughair, J. (1998). Pitfalls and opportunities in the use of extreme value theory in risk management. In Refenes, A., Moody, J., and Burgess, A., editors, *Decision Technologies for Computational Finance*, volume 2 of *Advances in Computational Management Science*, chapter 1, pages 3–12. Springer US.

Diebold, F. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.

Dupuis, D. J. and Tawn, J. A. (2001). Effects of mis-specification in bivariate extreme value problems. *Extremes*, 4(4):315–330.

Eckernkemper, T. (2017). Modeling systemic risk: Time-varying tail dependence when forecasting marginal expected shortfall. *Journal of Financial Econometrics*, 16(1):63–117.

Embrechts, P., Resnick, S. I., and Samorodnitsky, G. (1999). Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3:30–41.

Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics*, 20(3):339–350.

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

Hall, S., Swamy, P., and Tavlas, G. (2016). Time-varying coefficient models: A proposal for selecting the coefficient driver sets. *Macroeconomic Dynamics*, pages 1–17.

Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35(3):705–730.

Harris, R. and Mazibas, M. (2013). Dynamic hedge fund portfolio construction: A semi-parametric approach. *Journal of Banking & Finance*, 37(1):139–149.

Harvey, A. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Cambridge University Press, Cambridge, UK.

Harvey, A. and Chakravarty, T. (2008). Beta-t-(E)GARCH. *Discussion Paper, University of Cambridge*.

Hoga, Y. (2018). A structural break test for extremal dependence in β-mixing random vectors. *Biometrika*, 105(3):627–643.

Huser, R., Davison, A. C., and Genton, M. G. (2016). Likelihood estimators for multivariate extremes. *Extremes*, (19):79–103.

Joe, H. (2014). *Dependence Modeling with Copulas*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Chapman & Hall/CRC.

Kelly, B. and Jiang, H. (2014). Tail risk and asset prices. *The Review of Financial Studies*, 27(10):2841–2871.

Kritzman, M., Li, Y., Page, S., and Rigobon, R. (2011). Principal components as a measure of systemic risk. *Journal of Portfolio Management*, 37(4):112–126.

Leadbetter, M., Lindgren, G., and Rootzen, H. (1983). *Extremes and related properties of random sequences and processes*. Springer.

Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187.

Ledford, A. W. and Tawn, J. A. (1997). Modeling dependence within joint tail regions. *Journal of the Royal Statistical Society - Series B*, 49(2):475–499.

Longin, F. M. and Solnik, B. (2001). Extreme correlation of international equity markets. *Journal of Finance*, 56(2):649–676.

Mao, G. and Zhang, Z. (2018). Stochastic tail index model for high frequency financial data with bayesian analysis. *Journal of Econometrics*, 205:470–487.

Massacci, D. (2016). Tail risk dynamics in stock returns: Links to the macroeconomy and global markets connectedness. *Management Science*, pages 1–18.

McNeil, A. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3–4):271 – 300.

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

Nelsen, R. (1999). *An Introduction to Copulas*. Springer.

Newey, W. and McFadden, D. (1994). *Handbook of Econometrics*, volume 4, chapter Large sample estimation and hypothesis testing, pages 2111–2245. Elsevier.

Oh, D. H. and Patton, A. (2018). Time-varying systemic risk: Evidence from a dynamic copula model of cds spreads. *Journal of Business and Economic Statistics*, 36(2):181–195.

Pakel, C., Shephard, N., Sheppard, K., and Engle, R. (2018). Fitting vast dimensional time-varying covariance models. *NYU Working Paper No. FIN-08-009*.

Poon, S.-H., Rockinger, M., and Tawn, J. (2004). Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *Review of Financial Studies*, 17(2):581–610.

Quintos, C., Fan, Z., and Phillips, P. (2001). Structural change tests in tail behaviour and the asian crisis. *Review of Economic Studies*, 68(3):633–663.

Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. Springer.

Segers, J. (2012). Max-stable models for multivariate extremes. *REVSTAT – Statistical Journal*, 10(1):61–82.

Sibuya, M. (1959). Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics*, 11(2):195–210.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, (8):229–231.

Zhang, X. and Schwaab, B. (2017). Tail risk in government bond markets and ECB asset purchases. *Working paper*.

Zhao, Z. and Zhang, Z. (2018). Semiparametric dynamic max-copula model for multivariate time series. *Journal of Royal Statistical Society - Series B*, 80(2):409–432.

Zhao, Z., Zhang, Z., and Chen, R. (2018). Modeling maxima with autoregressive conditional Fréchet model. *Journal of Econometrics*, 207(2):325–351.

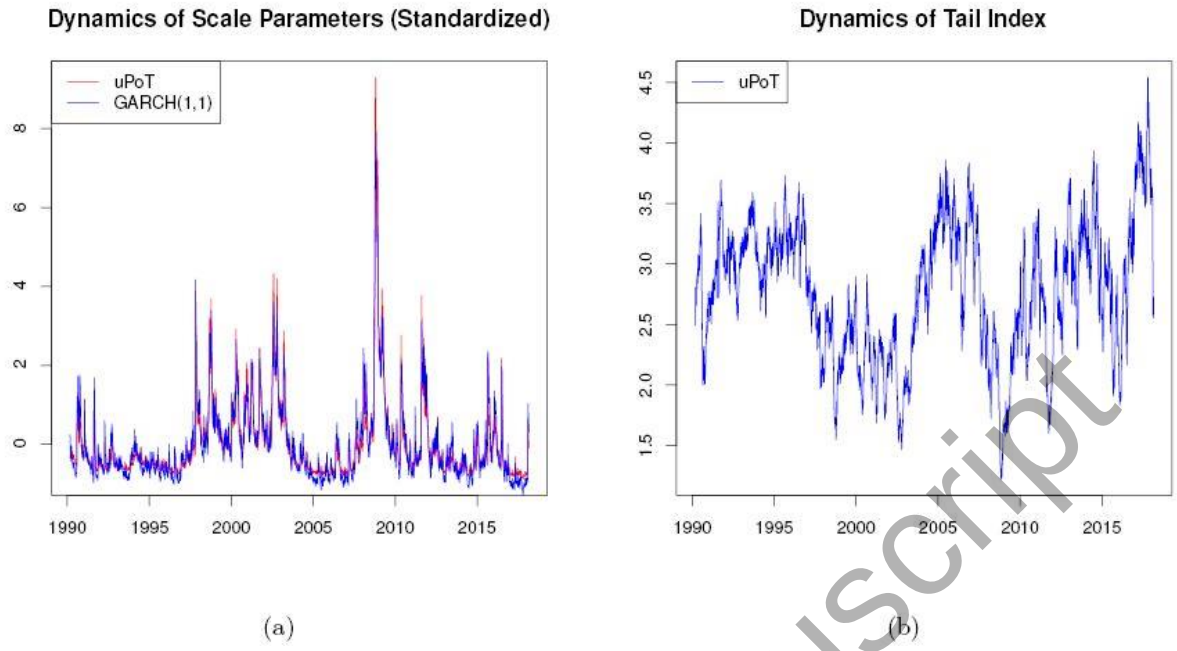

Fig. 1 *(a) Standardized average scale parameter* $\{\hat{\bar{\sigma}}_t\}$ *by dynamic uPoT and* $\{\tilde{\bar{\sigma}}_t\}$ *by GARCH(1,1). (b) Average tail index* $\{\hat{\bar{\xi}}_t\}$ *by dynamic uPoT.*

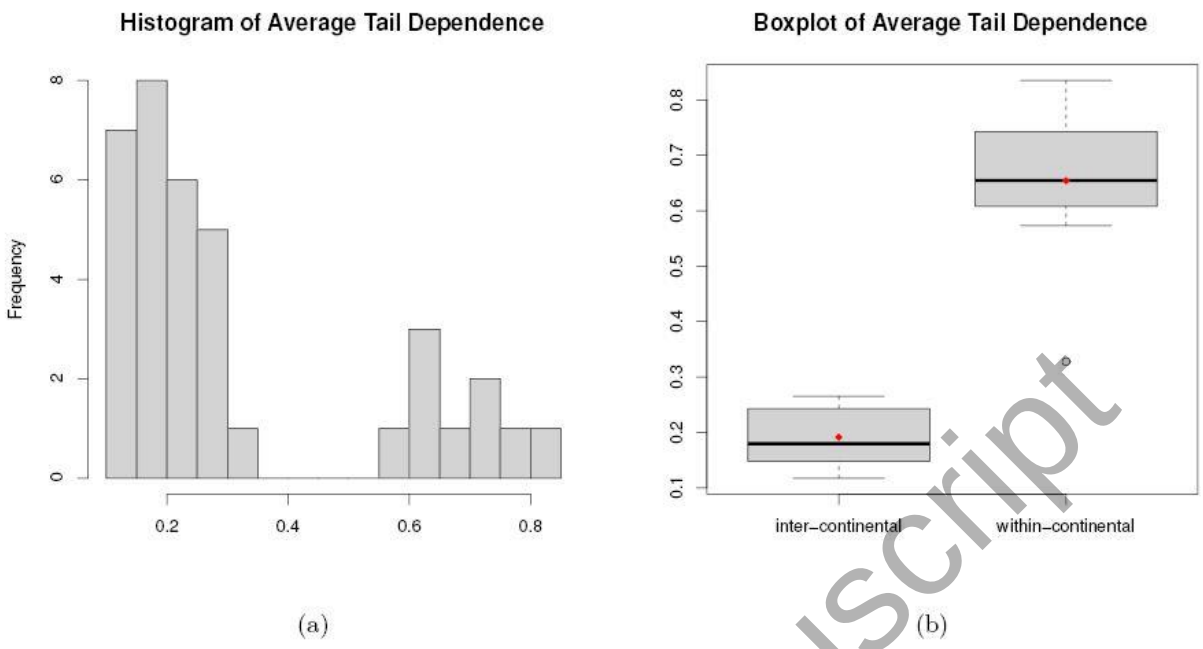

Fig. 2 *(a) Histogram of average tail dependence* $\{\hat{\bar{\lambda}}_k\}_{k=1}^{36}$*. (b) Boxplot of average tail dependence* $\{\hat{\bar{\lambda}}_k\}_{k=1}^{36}$ *by group. The red dot denotes group mean.*

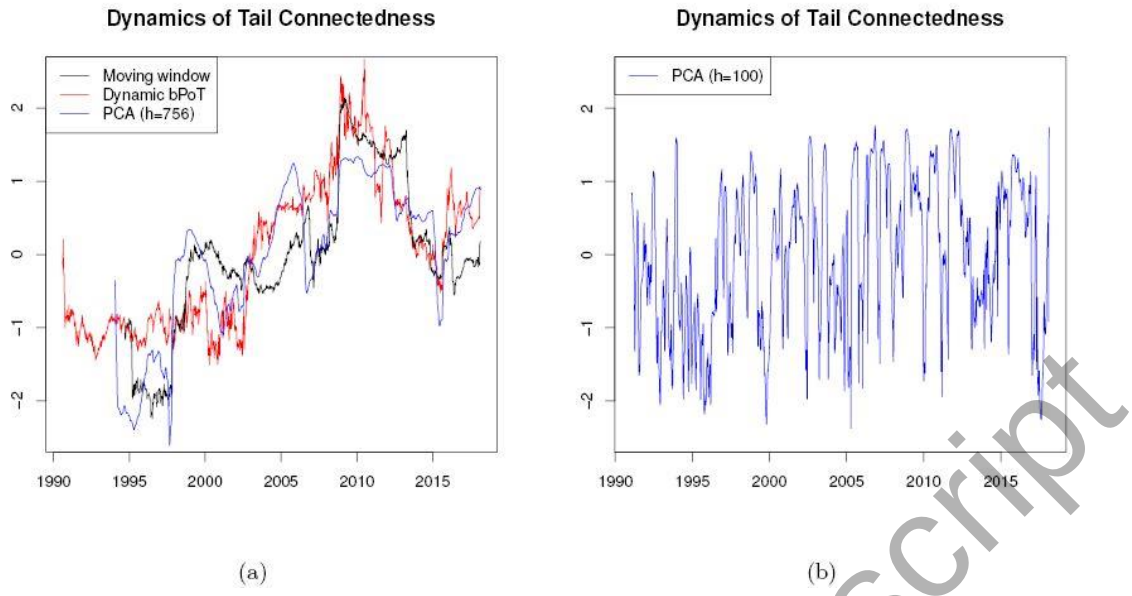

Fig. 3 *Standardized tail connectedness measures based on three different methods.*

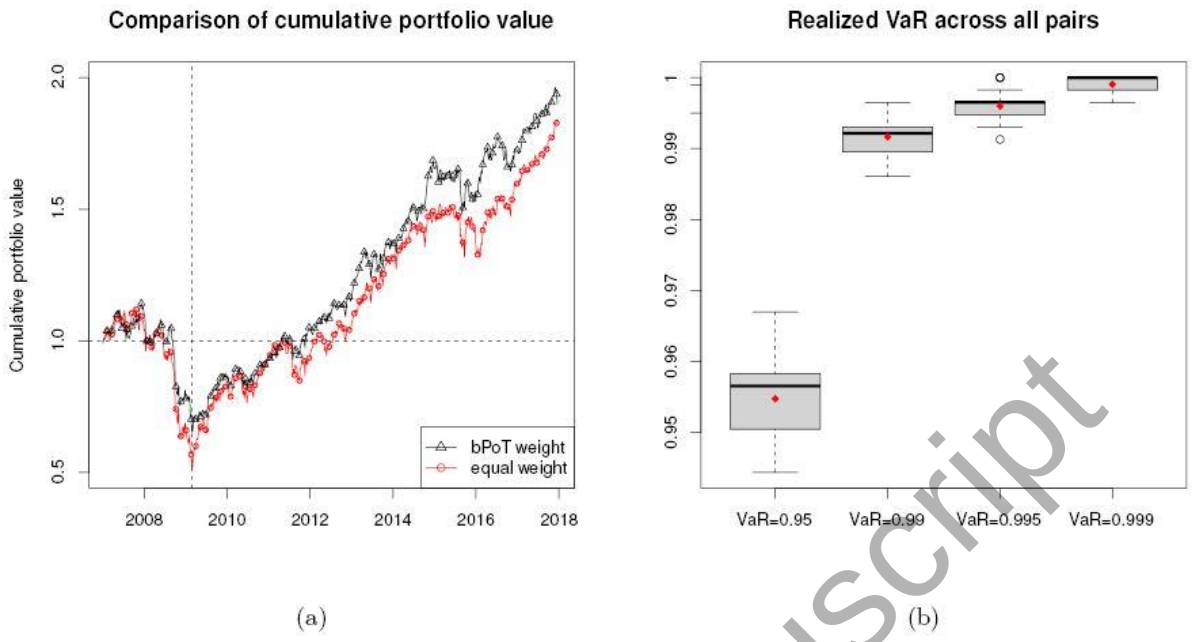

Fig. 4 *(a) Cumulative portfolio values over the period Jan 03, 2007 - Dec 29, 2017. The cumulative portfolio value on day t is calculated using* $\exp(\sum_{s=1}^{t} \hat{r}_s)$, *where* $\hat{r}_s$ *denotes the portfolio return on day s. (b) Boxplot of realized VaR level for 36 pairs of sector indices.*

Table 1 *Performance of full MLE and two-stage MLE under* $T = 3000, 6000, 9000.$ *S.D. stands for the sample standard deviation over 500 experiments.*

Full MLE		ψ_{01}	ψ_{11}	ψ_{21}	ϕ_{01}	ϕ_{11}	ϕ_{21}	β_0	β_1	β_2
sample size	true $\boldsymbol{\theta}$	0.033	0.816	0.0050	-0.017	0.949	0.112	0.010	0.990	0.100
T = 3000	Mean	0.032	0.815	0.0049	-0.014	0.948	0.106	0.020	0.981	0.105
	S.D.	0.012	0.060	0.0014	0.008	0.012	0.019	0.021	0.020	0.056
T = 6000	Mean	0.033	0.812	0.0050	-0.015	0.950	0.107	0.014	0.986	0.101
	S.D.	0.009	0.044	0.0010	0.006	0.008	0.013	0.011	0.010	0.030
T = 9000	Mean	0.033	0.814	0.0051	-0.015	0.950	0.107	0.011	0.989	0.099
	S.D.	0.007	0.037	0.0008	0.004	0.006	0.010	0.005	0.005	0.023
Two-stage MLE		ψ_{01}	ψ_{11}	ψ_{21}	ϕ_{01}	ϕ_{11}	ϕ_{21}	β_0	β_1	β_2
sample size	true $\boldsymbol{\theta}$	0.033	0.816	0.0050	-0.017	0.949	0.112	0.010	0.990	0.100
T = 3000	Mean	0.034	0.805	0.0050	-0.006	0.940	0.107	0.017	0.983	0.092
	S.D.	0.016	0.082	0.0020	0.024	0.032	0.041	0.019	0.018	0.045
T = 6000	Mean	0.035	0.803	0.0052	-0.013	0.947	0.107	0.014	0.986	0.096

Full MLE		ψ_{01}	ψ_{11}	ψ_{21}	ϕ_{01}	ϕ_{11}	ϕ_{21}	β_0	β_1	β_2
	S.D.	0.014	0.071	0.0014	0.015	0.018	0.027	0.011	0.011	0.029
T = 9000	Mean	0.034	0.808	0.0052	-0.014	0.946	0.111	0.011	0.989	0.096
	S.D.	0.012	0.065	0.0013	0.012	0.016	0.022	0.005	0.005	0.024

Table 2 *Average Pearson correlation and Spearman's rho between the recovered process* $\{(\hat{\sigma}_{t1}, \hat{\xi}_{t1}, \hat{\alpha}_t)\}_{t=1}^{T}$ *and the true process* $\{(\sigma_{t1}, \xi_{t1}, \alpha_t)\}_{t=1}^{T}$ *over 500 experiments. The standard deviation is reported in the brackets.*

	Pearson correlation			Spearman's rho		
Full MLE	$\rho(\sigma_{t1}, \hat{\sigma}_{t1})$	$\rho(\xi_{t1}, \hat{\xi}_{t1})$	$\rho(\alpha_t, \hat{\alpha}_t)$	$\rho(\sigma_{t1}, \hat{\sigma}_{t1})$	$\rho(\xi_{t1}, \hat{\xi}_{t1})$	$\rho(\alpha_t, \hat{\alpha}_t)$
T = 3000	0.983 (0.024)	0.994 (0.012)	0.921 (0.070)	0.993 (0.013)	0.979 (0.029)	0.919 (0.073)
T = 6000	0.992 (0.012)	0.997 (0.004)	0.957 (0.036)	0.997 (0.004)	0.990 (0.015)	0.956 (0.037)
T = 9000	0.994 (0.009)	0.998 (0.002)	0.973 (0.018)	0.998 (0.002)	0.993 (0.011)	0.971 (0.020)
	Pearson correlation			Spearman's rho		
Two-stage	$\rho(\sigma_{t1}, \hat{\sigma}_{t1})$	$\rho(\xi_{t1}, \hat{\xi}_{t1})$	$\rho(\alpha_t, \hat{\alpha}_t)$	$\rho(\sigma_{t1}, \hat{\sigma}_{t1})$	$\rho(\xi_{t1}, \hat{\xi}_{t1})$	$\rho(\alpha_t, \hat{\alpha}_t)$
T = 3000	0.970 (0.051)	0.967 (0.034)	0.894 (0.075)	0.967 (0.055)	0.959 (0.041)	0.893 (0.080)
T = 6000	0.987 (0.024)	0.978 (0.025)	0.939 (0.044)	0.986 (0.027)	0.973 (0.031)	0.939 (0.044)
T = 9000	0.992 (0.018)	0.982 (0.021)	0.961 (0.021)	0.991 (0.019)	0.977 (0.025)	0.960 (0.022)

Table 3 *Estimation results of the dynamic uPoT model for each stock index. Standard errors are in reported parentheses below parameter estimates.*

	Scale parameter σ_t			Shape parameter ξ_t			Correlation	
	ψ_0	ψ_1	ψ_2 $(\times 10^{-2})$	ϕ_0	ϕ_1	ϕ_2	$\rho(\hat{\sigma}_{ti}, \tilde{\sigma}_{ti})$	$\rho(\Delta\hat{\sigma}_{ti}, \Delta\tilde{\sigma}_{ti})$
DJI	0.033	0.816	0.209	-0.017	0.949	0.113	0.948	0.924
	(0.015)	(0.078)	(0.118)	(0.006)	(0.012)	(0.026)	-	-
SP500	0.022	0.891	0.101	-0.019	0.940	0.132	0.954	0.912
	(0.020)	(0.099)	(0.097)	(0.006)	(0.013)	(0.029)	-	-
NASDAQ	0.038	0.899	0.245	-0.003	0.947	0.113	0.979	0.942
	(0.040)	(0.104)	(0.219)	(0.005)	(0.017)	(0.033)	-	-
FTSE	0.060	0.727	0.179	-0.023	0.933	0.159	0.897	0.895
	(0.043)	(0.192)	(0.175)	(0.008)	(0.018)	(0.042)	-	-
CAC40	0.048	0.864	0.190	-0.006	0.929	0.155	0.971	0.949
	(0.030)	(0.082)	(0.137)	(0.007)	(0.017)	(0.035)	-	-
DAX	0.038	0.907	0.193	0.002	0.915	0.175	0.978	0.951
	(0.044)	(0.103)	(0.189)	(0.009)	(0.021)	(0.038)	-	-
ES50	0.054	0.829	0.340	-0.009	0.947	0.120	0.960	0.951
	(0.023)	(0.067)	(0.159)	(0.005)	(0.013)	(0.030)	-	-

	Scale parameter σ_t			Shape parameter ξ_t			Correlation	
HSI	0.004	0.990	0.019	-0.006	0.942	0.127	0.774	0.831
	(0.002)	(0.005)	(0.012)	(0.006)	(0.010)	(0.020)	-	-
N225	0.055	0.896	0.029	0.007	0.901	0.201	0.926	0.902
	(0.083)	(0.157)	(0.073)	(0.010)	(0.015)	(0.030)	-	-

Table 4 *The pairs of stock indices with statistically significant time-varying tail coefficient, its average tail coefficient* $\hat{\bar{\alpha}}_k = \sum_{t=1}^{6202} \hat{\alpha}_{tk} / 6202$ *, median tail coefficient* $\hat{\underline{\alpha}}_k = \text{median}\{\hat{\alpha}_{tk}\}_{t=1}^{6202}$ *and* χ^2 *-value for the likelihood ratio test.*

Pair of stock indices	$\hat{\bar{\alpha}}_k$	$\hat{\underline{\alpha}}_k$	χ^2 -value	Pair of stock indices	$\hat{\bar{\alpha}}_k$	$\hat{\underline{\alpha}}_k$	χ^2 -value
DJI v.s. SP500	4.71	4.68	73.5	FTSE v.s. ES50	2.17	2.03	40.4
CAC40 v.s. ES50	4.12	4.62	97.5	FTSE v.s. DAX	2.01	1.92	44.4
DAX v.s. ES50	3.29	3.26	74.3	HSI v.s. N225	1.36	1.34	30.1
SP500 v.s. NASDAQ	2.96	2.79	57.6	SP500 v.s. N225	1.23	1.22	16.4
CAC40 v.s. DAX	2.66	2.64	82.5	DJI v.s. N225	1.23	1.21	16.3
DJI v.s. NASDAQ	2.32	2.21	101.9	NASDAQ v.s. N225	1.21	1.19	14.2
CAC40 v.s. FTSE	2.29	2.17	55.4				

Table 5 *Selected estimation results for pairs of stock indices with statistically significant time-varying tail coefficients. Standard errors are reported in parentheses below parameter estimates.*

Pair of stock indices (Top 3)	β_0	β_1	β_2	$\hat{\bar{\alpha}}_k$	$\hat{\underline{\alpha}}_k$
DJI v.s. SP500	0.0056	0.9955	0.0592	4.71	4.68
	(0.0044)	(0.0038)	(0.0222)		
CAC40 v.s. ES50	0.0004	0.9996	0.0359	4.12	4.62
	(0.0003)	(0.0004)	(0.0111)		
DAX v.s. ES50	0.0006	0.9993	0.0614	3.29	3.26
	(0.0004)	(0.0006)	(0.0158)		
Pair of stock indices (Bottom 3)	β_0	β_1	β_2	$\hat{\bar{\alpha}}_k$	$\hat{\underline{\alpha}}_k$
DJI v.s. N225	-0.0239	0.9844	0.3235	1.23	1.22
	(0.0161)	(0.0102)	(0.1287)		
SP500 v.s. N225	-0.0190	0.9873	0.2863	1.23	1.21
	(0.0150)	(0.0097)	(0.1239)		
NASDAQ v.s. N225	-0.0050	0.9969	0.1620	1.21	1.19
	(0.0180)	(0.0112)	(0.1725)		

Table 6 *Selected estimation results for pairs of indices with statistically insignificant time-varying tail coefficients. Standard errors are reported in parentheses below parameter estimates.*

Pair of stocks (Top 3)	HSI v.s. SP500	HSI v.s. ES50	HSI v.s. FTSE
$\hat{\alpha}$	1.259	1.255	1.253
	(0.022)	(0.022)	(0.021)
Pair of stocks (Bottom 3)	DAX v.s. SP500	DAX v.s. NASDAQ	DAX v.s. DJI
$\hat{\alpha}$	1.096	1.095	1.094
	(0.015)	(0.015)	(0.015)

Table 7 *Summary statistics of weekly returns over the period Jan 03, 2007 - Dec 29, 2017. "AR" stands for annualized return and "SR" stands for Sharpe ratio. All numbers, except SR, are in percentage.*

Scenario	Portfolio	AR	SD	SR	VaR 95%	VaR 99.5%	ES 95%	ES 99.5%
Overall	bPoT-weight	6.04	1.80	0.065	2.67	7.81	4.55	10.23
	equal-weight	5.71	2.32	0.047	3.46	10.09	5.83	12.18
	S&P500	6.00	2.52	0.046	4.43	9.17	6.30	13.17
Crisis	bPoT-weight	-20.37	2.65	-0.148	5.15	10.73	8.40	12.17
	equal-weight	-31.99	3.36	-0.183	5.19	12.87	10.01	15.62
	S&P500	-35.13	3.39	-0.199	6.02	12.61	9.27	14.91
Post-crisis	bPoT-weight	12.52	1.50	0.161	2.30	4.36	3.13	4.56
	equal-weight	14.97	1.94	0.148	2.83	5.85	4.22	6.25
	S&P500	16.09	2.21	0.140	3.42	5.72	4.88	8.87