Biometrika (20XX), **XXX**, X, *pp.* 1–14
Printed in Great Britain

A structural break test for extremal dependence in β-mixing random vectors

BY Y. HOGA

Faculty of Economics and Business Administration, Universität Duisburg-Essen, Universitätsstraße 12, D-45117 Essen, Germany,

yannick.hoga@vwl.uni-due.de

SUMMARY

We derive a structural break test for extremal dependence in β-mixing, possibly high-dimensional random vectors with either asymptotically dependent or asymptotically independent components. Existing tests require serially independent observations with asymptotically dependent components. To avoid estimating a long-run variance, we use self-normalization, which obviates the need to estimate the coefficient of tail dependence when components are asymptotically independent. Simulations show favourable empirical size and power of the test, which we apply to S&P 500 and DAX log-returns. We find evidence for one break in the coefficient of tail dependence of the upper and lower joint tail at the beginning of the financial crisis of 2007–8, leading to more extremal co-movement.

Some key words: β-mixing; Extremal dependence; Self-normalization, Structural break test.

1. MOTIVATION

The study of extremal properties of time series has received considerable attention recently, not least because of the financial crisis of 2007–8. In financial risk management, for instance, one may be interested in the value-at-risk, i.e., a small quantile of the return distribution of risky assets, or the expected shortfall, i.e., the average loss beyond the value-at-risk. Beyond these univariate distributional quantities, bivariate extremal dependence between the components of a random vector $(X_i, Y_i)^{\mathrm{T}}$ is also important. There are many ways to measure extremal dependence, for instance the tail dependence coefficient

$$\lambda = \lim_{s\to\infty} \lambda_s = \lim_{s\to\infty} \mathrm{pr}\left\{X_i > F_X^{\leftarrow}(1-1/s) \mid Y_i > F_Y^{\leftarrow}(1-1/s)\right\},$$

where it is assumed that the limit exists, and F_X and F_Y denote the distribution functions of X_i and Y_i. The tail dependence coefficient dates back at least to Sibuya (1960). If $\lambda > 0$, we have asymptotic dependence, and if $\lambda = 0$ there is asymptotic independence. Davis & Mikosch (2009) and Davis et al. (2013) introduce the extremogram and pre-asymptotic extremogram as generalizations of λ and λ_s, respectively.

A popular framework for assessing tail dependence in the case of asymptotic independence, i.e., $\lambda = 0$, was developed by Ledford & Tawn (1996, 1997). Transforming the continuous X_i and Y_i to standard Fréchet marginals $\widetilde{X}_i = -1/\log F_X(X_i)$ and $\widetilde{Y}_i = -1/\log F_Y(Y_i)$, they model the joint tail as

$$\mathrm{pr}(\widetilde{X}_i > z, \widetilde{Y}_i > z) \sim L(z)\, \mathrm{pr}(\widetilde{Y}_i > z)^{1/\eta},$$

where $\eta \in (0, 1]$ is the coefficient of tail dependence given tail independence and $L(\cdot)$ is a slowly-varying function. If $\eta = 1$ and $\lim_{z\to\infty} L(z) = c \in (0, 1]$, then $\lambda = \lim_{z\to\infty} \mathrm{pr}(\widetilde{X}_i > z \mid \widetilde{Y}_i > z) = c$. If $\eta < 1$, then $\lambda = 0$ and η can be viewed as quantifying the speed of convergence of the conditional probability to the limit $\lambda = 0$. Here, $\eta > 1/2$, $\eta = 1/2$ and $\eta < 1/2$ correspond to positive association, independence and negative association in the tail, respectively. Coles et al. (1999) and Poon et al. (2004) argue that the pair $(\chi = \lambda, \overline{\chi} = 2\eta - 1)$ offers a concise description of the extremal dependence in $(X_i, Y_i)^{\mathrm{T}}$.

The purpose of the present paper is to derive a structural break test for the extremal dependence between the components of the random vectors $(X_1, Y_1)^{\mathrm{T}}, \ldots, (X_n, Y_n)^{\mathrm{T}}$, while the margins are assumed to be identically distributed through time. More precisely, we consider

$$p_{n,i}^{(\eta)} = \big(n/k_n\big)^{1/\eta} \operatorname{pr}(X_i > b_{x,n}, Y_i > b_{y,n}), \tag{1}$$

where $b_{x,n}$ and $b_{y,n}$ are the $(1 - k_n/n)$-quantiles of X_i and Y_i, $k_n = o(n)$, and $\eta \in (0, 1]$ is the coefficient of tail dependence given tail independence in Ledford & Tawn's framework.

The use of $p_{n,i} = p_{n,i}^{(1)}$ has several advantages: it is convenient to calculate, invariant under transformations of the marginal distributions, nonparametric and joint tail dependence model-free. It does not offer a complete description of the extremal dependence structure, as it is merely a scalar, but, noting that $p_{n,i}$ is essentially $\operatorname{pr}(X_i > b_{x,n} \mid Y_i > b_{y,n})$, it may be interpreted as a conditional probability. Thus, if $Y_i = X_{i-h}$ and we are interested in serial extremal dependence, it is a useful scalar, as in 'financial applications [...] one is often interested in the persistence of a shock (an extremal event on the stock market say) at future instants of time' (Davis et al., 2012, p. 143). When Y_i is not a lagged X_i but, say, the return of another stock market, it may be used as a measure of how contagious shocks are in the financial system (Bae et al., 2003; Poon et al., 2004), in which case it is again an interesting summary measure. Poon et al. (2004, p. 597) identify the limit of $p_{n,i}$ as a 'true measure of systemic risk'. Finally, $p_{n,i}$ may be scaled to give $p_{n,i}^{(\eta)}$ so that its limit is revealing even when $\eta < 1$, and hence $\lambda = 0$ is uninformative.

While break detection in general measures of dependence, e.g., the covariance, has been studied extensively (e.g., Aue et al., 2009; Wied et al., 2012), only Bücher et al. (2015) propose a structural break test for the tail dependence coefficient. Tail dependence is crucial in many financial applications, where dependencies may only be inadequately captured by traditional measures (Embrechts et al., 2002; Bae et al., 2003; Poon et al., 2004). For instance, the Pearson correlation neither distinguishes large/small nor positive/negative returns and hence frequently underestimates the risk of joint extreme events.

Bücher et al. (2015) require temporally independent data, which is implausible in financial contexts. We address this by allowing for β-mixing data. To obviate the need to estimate a long-run variance in our test statistic, which is typically complicated by the presence of serial dependence, we use self-normalization (Shao & Zhang, 2010).

As a second extension of Bücher et al. (2015), who assume asymptotically dependent data with $\eta = 1$, we also allow for weaker forms of dependence between X_i and Y_i, where $\eta < 1$. This substantially widens applicability. For instance, as shown by Poon et al. (2003, 2004), different strengths of linkages corresponding to $\eta = 1$ and $\eta < 1$ may be found between stock indices. Stock index returns in the U.K., Germany and France appear to have strong extremal dependence, with $\eta = 1$. However, the returns of stock markets in Europe, the U.S.A. and Japan exhibit $\eta < 1$. By virtue of self-normalization it turns out that our test may be applied without any knowledge or need for estimation of η.

Our third main contribution is to study asymptotics under local alternatives and alternatives with a change in η. The former alternatives are interesting, because the local power of change-point tests in extreme-value settings can be different from standard $n^{-1/2}$-results in, e.g., Wied et al. (2012), and typically depends on k_n (Hoga, 2017c). The latter alternatives are relevant in empirical applications, where preliminary evidence suggests possible breaks in η for stock index returns (Poon et al., 2003, 2004).

Finally, we consider testing for multiple changes, employing a variant of self-normalization proposed by Zhang & Lavitas (2018+). Unlike the extensions suggested by Shao & Zhang (2010) to deal with more than one break, their approach does not require the number of change points to be pre-specified and has a computational burden that does not increase in the number of possible change-points.

2. MAIN RESULTS

2·1. *Preliminaries*

Consider a bivariate $\mathbb{R}^2$-valued stochastic process $\big\{V_{n,i} = (X_i, Y_i)^{\mathrm{T}}\big\}_{n \in \mathbb{N}, i=1,\ldots,n}$. Let $Z_i \in \{X_i, Y_i\}$ be a generic component of $V_{n,i}$ and denote by F_Z its time-invariant distribution function. Strictly speaking, we should write $V_{n,i} = (X_{n,i}, Y_{n,i})^{\mathrm{T}}$. However, for notational convenience and to emphasize the time-invariance of the marginal distributions, we shall suppress the array notation in the marginals.

Even under the alternative of a change in the extremal dependence, we assume identically distributed marginals. Bücher et al. (2015) attempt to weaken that assumption in their independent setting by allowing for a one-time break in the marginal distributions. Yet in this case, the limiting distribution of their test statistic is not tractable, as it depends on the unknown breakpoint of the marginal distributions. We leave the task of allowing for non-identically distributed marginals in our test for future study.

Denote by $Z_{(1)} \geq \cdots \geq Z_{(n)}$ the order statistics of a sample $Z_1, \ldots, Z_n$. As we are interested in tail dependence, we need some intermediate sequence $k_n \in \mathbb{N}$ with $k_n \leq n-1$, such that $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. This sequence controls the number of large observations used in the estimation of the extremal dependence of X_i and Y_i, and hence specifies where the tail begins.

Since $k_n/n \to 0$, the quantity $p_{n,i}^{(\eta)}$ in (1) is a measure of tail dependence. A natural estimator of $p_{n,i}^{(\eta)}$ is its empirical analogue

$$\widehat{p}_n = \left(\frac{n}{k_n}\right)^{1/\eta} \frac{1}{n} \sum_{i=1}^{n} I_{\left(X_i > X_{(k_n+1)},\, Y_i > Y_{(k_n+1)}\right)}, \tag{2}$$

where $I_{(\cdot)}$ denotes the indicator function. We shall see that there is no need to estimate η for our test to work. Introduce $c_n = c_n(\eta) = n(k_n/n)^{1/\eta}$, which will turn out to be the squared convergence rate of $\widehat{p}_n$.

If there is doubt whether extremal dependence in $(X_1, Y_1)^{\mathrm{T}}, \ldots, (X_n, Y_n)^{\mathrm{T}}$ is constant over time, one may wish to test the hypothesis

$$\begin{aligned} \mathcal{H}_0^{(\eta)}: &\quad p_{n,i}^{(\eta)} = p_{n,1}^{(\eta)} + o(c_n^{-1/2}), \\ \mathcal{H}_1^{(\eta)}: &\quad p_{n,i}^{(\eta)} = p_{n,1}^{(\eta)} + o(c_n^{-1/2}) + Mc_n^{-1/2} I_{(i > \lfloor nt^* \rfloor)}, \end{aligned} \qquad i = 1, \ldots, n, \tag{3}$$

where $t^* \in (0,1)$ denotes the breakpoint, $M \neq 0$ the magnitude of the break and the $o(c_n^{-1/2})$-terms are uniform in i. In the leading case $\eta = 1$, we simply write $\mathcal{H}_0$, $\mathcal{H}_1$, $p_{n,i}$ instead of $\mathcal{H}_0^{(1)}$, $\mathcal{H}_1^{(1)}$, $p_{n,i}^{(1)}$.

At first sight it may appear that the null to be tested depends on n. However, as (3) is an asymptotic relation, this is not the case. The null to be tested depends on k_n, which is restricted by Assumptions 2 and 3 below. The faster k_n is allowed to grow, the smaller the pre-asymptotic differences, $Mc_n^{-1/2}$, we can detect. This means that the more observations we can use in estimation, the easier it is to detect changes in the extremal dependence. Also, the larger η, i.e., the more heavily dependent X_i and Y_i, the smaller the pre-asymptotic differences, $Mc_n^{-1/2}$, we can detect. We consider alternatives in a $c_n^{-1/2}$-neighbourhood, whereas usually $n^{-1/2}$-neighbourhoods are studied. This is because the estimator $\widehat{p}_n$ is essentially $c_n^{1/2}$-consistent and not $n^{1/2}$-consistent. Since for larger η more observations will lie in the joint tail, it is quite intuitive that the convergence rate $c_n^{1/2}$ of $\widehat{p}_n$ increases in η.

The case $\eta = 1$ is dealt with for serially independent $\{(X_i, Y_i)^{\mathrm{T}}\}$ by Bücher et al. (2015). They assume $p_{n,i} = (n/k_n)\,\mathrm{pr}(X_i > b_{x,n}, Y_i > b_{y,n}) \to \lambda_i > 0$ as $n \to \infty$, and test

$$\begin{aligned} \mathcal{H}_0^{\lambda}: &\quad \lambda_1 = \cdots = \lambda_n, \\ \mathcal{H}_1^{\lambda}: &\quad \lambda_1 = \cdots = \lambda_{\lfloor nt^* \rfloor} \neq \lambda_{\lfloor nt^* \rfloor + 1} = \cdots = \lambda_n, \qquad t^* \in (0,1). \end{aligned}$$

There are several reasons why we prefer to test $\mathcal{H}_0$ rather than $\mathcal{H}_0^{\lambda}$. If $\lim_{n\to\infty} p_{n,1} = \lambda$, then under both $\mathcal{H}_0$ and $\mathcal{H}_1$ we have $\lim_{n\to\infty} p_{n,i} = \lambda$. Hence, $\mathcal{H}_1$ may be termed a local alternative, as both $\mathcal{H}_0$ and $\mathcal{H}_1$ are covered under $\mathcal{H}_0^{\lambda}$. Yet Theorem 2 below shows that our test has non-trivial power for $\mathcal{H}_1$. Thus, to test $\mathcal{H}_0^{\lambda}$ Bücher et al. (2015, Ass. 3.1 & 3.2) impose the additional second-order type assumption

$$k_n^{1/2} \left(p_{n,i} - \lambda_i\right) \longrightarrow 0, \quad n \to \infty, \tag{4}$$

uniformly in $i = 1, \ldots, n$, which prohibits pre-asymptotic fluctuations as under $\mathcal{H}_1$. Theorem 2 shows that the test by Bücher et al. (2015) is not only a test of $\mathcal{H}_0^{\lambda}$, but also implicitly of (4), because a violation of (4), even without an accompanying change in λ_i, ultimately leads to a rejection.

For illustration, assume that (4) holds for $i = 1, \dots, \lfloor nt^* \rfloor$, yet for $i = \lfloor nt^* \rfloor + 1, \dots, n$ we have

$$k_n^{1/2}(p_{n,i} - \lambda_i) \longrightarrow M \neq 0, \quad n \to \infty.$$

Also assume $\lambda_1 = \cdots = \lambda_n$. Then we are under $\mathcal{H}_0^\lambda$, yet $k_n^{1/2}(p_{n,1} - p_{n,n}) \to M \neq 0$ as $n \to \infty$, which is only covered by $\mathcal{H}_1$. Theorem 2 shows that, as the size of $|M|$ grows, $\mathcal{H}_0^\lambda$ will eventually be rejected even though $\mathcal{H}_0^\lambda$ is true. So one reason we prefer a test of the pair $\mathcal{H}_0, \mathcal{H}_1$ is that they are more explicit about what exactly is being tested.

Assumption (4) also presupposes a lot of knowledge about λ_i, which is not even known to be constant in the sample. Finally, (4) may be hard to verify. For instance, Davis & Mikosch (2009, Sec. 4.1) could not verify it for $(X_i, Y_i = X_{i-h})^{\mathrm{T}}$ with $\{X_i\}$ a GARCH$(1,1)$-process.

The serial dependence concept we use is that of β-mixing. A possibly triangular sequence of random vectors $\{V_{n,i}\}_{n\in\mathbb{N}, i=1,\dots,n}$ is β-mixing if

$$\beta_n(l) = \sup_{m\in\{1,\dots,n-l-1\}} E\left\{ \sup_{A\in\mathcal{F}_{n,m+l+1}^n} \left| \mathrm{pr}(A \mid \mathcal{F}_{n,1}^m) - \mathrm{pr}(A) \right| \right\} \longrightarrow 0, \quad l \to \infty, \tag{5}$$

where $\mathcal{F}_{n,l}^m$ is the σ-algebra generated by $\{V_{n,l}, \dots, V_{n,m}\}$. We allow for a triangular array to also cover the alternative $\mathcal{H}_1^{(\eta)}$ in our development. If $V_{n,i} = V_i$ is a sequence of β-mixing random vectors, then array β-mixing in the sense of (5) follows.

As a model for the joint tail of $(X_i, Y_i)^{\mathrm{T}}$, we impose the Ledford–Tawn model

$$\overline{F}_i(x,y) = \mathrm{pr}\left(\widetilde{X}_i > x, \widetilde{Y}_i > y\right) = \mathcal{L}_i(x,y)(xy)^{-1/(2\eta)}, \qquad \eta \in (0,1],$$

where $\mathcal{L}_i(\cdot,\cdot)$ is a bivariate slowly-varying function, i.e., there exists a function $g_i(\cdot,\cdot)$ such that for all $x, y, c > 0$ it holds that $g_i(cx, cy) = g_i(x,y)$ and

$$g_i(x,y) = \lim_{r\to\infty} \frac{\mathcal{L}_i(rx, ry)}{\mathcal{L}_i(r,r)}. \tag{6}$$

We can now state our main assumptions.

Assumption 1. The marginals of $V_{n,i}$ have identical continuous distribution functions. The joint tail of $V_{n,i}$ is governed by the Ledford–Tawn model, where (6) holds uniformly on $\{(x,y)^{\mathrm{T}} \in (0,\infty)^2 : x^2 + y^2 = 1\}$ with $\limsup_{r\to\infty} \mathcal{L}_i(r,r) \leq C < \infty$ for C independent of $i \in \mathbb{N}$.

Assumption 2. The process $\{V_{n,i}\}_{n\in\mathbb{N}, i=1,\dots,n}$ is β-mixing with mixing coefficients $\beta_n(\cdot)$, such that

$$\lim_{n\to\infty} \left\{ \frac{n}{r_n} \beta_n(l_n) + r_n c_n^{-1/2}(\eta) \right\} = 0$$

for integer sequences $\{l_n\}_{n\in\mathbb{N}}$, $\{r_n\}_{n\in\mathbb{N}}$ tending to infinity with $l_n = o(r_n)$ and $r_n = o(n)$.

Assumption 3. The following limits are uniform in $j \geq 0$:

$$\lim_{n\to\infty} \frac{1}{r_n} \left(\frac{n}{k_n}\right)^{1/\eta} \mathrm{var}\left\{ \sum_{i=j+1}^{j+r_n} I_{(X_i > b_{x,n}, Y_i > b_{y,n})} \right\} = \sigma^2 \in (0,\infty), \tag{7}$$

$$\lim_{n\to\infty} \frac{1}{r_n} \left(\frac{n}{k_n}\right)^{1/\eta} \mathrm{var}\left\{ \sum_{i=j+1}^{j+l_n} I_{(X_i > b_{x,n}, Y_i > b_{y,n})} \right\} = 0.$$

Remark 1. By a probability integral transform-type argument, Assumption 1 allows us to prove the main results for $V_{n,i}$ with standard Fréchet-distributed marginals. The Ledford–Tawn model is a mild assumption on the joint tail, as '[a]ll bivariate extreme value dependence structures and the great majority of dependence structures that exist within the copula literature have extremal dependence structures that can be represented in the form' (Ramos & Ledford, 2009, p. 221). Assumption 2 allows for the application of a standard big block/small block argument. The small blocks of length l_n are asymptotically negligible and the big blocks of length r_n converge to some well-defined limit by virtue of a standard functional central limit theorem. Similar mixing conditions have been used by, e.g., Drees & Rootzén (2010) and Hoga (2017a). Assumption 3 facilitates the application of functional central limit theory for the big blocks and small blocks. Proposition 1 below provides more easily verified sufficient conditions for Assumption 3.

Remark 2. Suppose that geometric β-mixing holds, i.e., $\beta_n(l_n) = \mathcal{O}(K^{l_n})$ for some $K \in (0,1)$, which is satisfied for many linear and non-linear processes (Liebscher, 2005; Meitz & Saikkonen, 2008). Then, for $l_n = \lceil -2\log n/\log K \rceil$ it is easy to show that $(n/r_n)\beta_n(l_n) = o(1)$ for any sequence $r_n \to \infty$. So if $k_n \sim an^b$ for $a > 0$, Assumption 2 is compatible with $b \in (1-\eta, 1)$, which is also required for the convergence rate $c_n^{1/2}(\eta) = n^{1/2}(k_n/n)^{1/(2\eta)}$ of $\widehat{p}_n$ to tend to infinity. Thus, Assumption 2 is not restrictive.

The next proposition gives easier-to-verify sufficient conditions for Assumption 3 to hold.

PROPOSITION 1. *Suppose that $V_{n,i} = V_i$ is a strictly stationary sequence of random vectors and that for all $m \in \mathbb{N}_0$*

$$\left(\frac{n}{k_n}\right)^{1/\eta} \mathrm{pr}\left(X_1 > b_{x,n}, Y_1 > b_{y,n}, X_{1+m} > b_{x,n}, Y_{1+m} > b_{y,n}\right) \longrightarrow c_m, \quad n \to \infty, \tag{8}$$

where $c_0 > 0$, and

$$\lim_{h\to\infty} \limsup_{n\to\infty} \left(\frac{n}{k_n}\right)^{1/\eta} \sum_{m=h+1}^{r_n} \mathrm{pr}\left(X_1 > b_{x,n}, Y_1 > b_{y,n}, X_{1+m} > b_{x,n}, Y_{1+m} > b_{y,n}\right) = 0. \tag{9}$$

Then, Assumption 3 is met with $\sigma^2 = c_0 + 2\sum_{m=1}^{\infty} c_m > 0$ if $\lim_{n\to\infty} r_n(k_n/n)^{1/\eta} = 0$.

Remark 3. An almost trivial, yet quite useful sufficient condition for (9) is

$$\lim_{h\to\infty} \limsup_{n\to\infty} \left(\frac{n}{k_n}\right)^{1/\eta} \sum_{m=h+1}^{r_n} \mathrm{pr}\left(X_1 > b_{x,n}, X_{1+m} > b_{y,n}\right) = 0. \tag{10}$$

Davis & Mikosch (2009) verify (10) for generalized autoregressive conditional heteroscedasticity and stochastic volatility models. Multivariate regular variation is sufficient for (8) in the serial extremal dependence case where Y_i is a lagged X_i, i.e., $Y_i = X_{i-h}$. Fasen et al. (2010) show that multivariate regular variation holds for many popular time series models.

2·2. *Results under the null and under the alternative*

For $0 \le s < t \le 1$, we define pseudo-subsample estimates of $p_{n,i}^{(\eta)}$ via

$$\widehat{p}_n(s,t) = \widehat{p}_{n,k_n}^{(\eta)}(s,t) = \left(\frac{n}{k_n}\right)^{1/\eta} \frac{1}{n(t-s)} \sum_{i=\lfloor ns\rfloor+1}^{\lfloor nt\rfloor} I_{\left(X_i > X_{(k_n+1)}, Y_i > Y_{(k_n+1)}\right)}. \tag{11}$$

We call this a pseudo-estimate, because η is typically unknown. In change-point analysis one frequently compares subsample estimates à la

$$G_n(t) = c_n^{1/2} t(1-t)\left\{\widehat{p}_n(0,t) - \widehat{p}_n(t,1)\right\}, \qquad t \in [0,1], \tag{12}$$

and rejects $\mathcal{H}_0^{(\eta)}$ if some test statistic based on $G_n(t)$, e.g., $\mathcal{T}_n = \widehat{\sigma}^{-2}\int_0^1 G_n^2(t)\mathrm{d}t$, is too large. However, estimates $\widehat{\sigma}^2$ of the long-run variance of $\widehat{p}_n(0,1)$ may lead to nonmonotonic power, where power of change-point tests does not increase the more extreme the alternative (Vogelsang, 1997, 1999). Hence, it is desirable to obviate the need to estimate the long-run variance. This is conveniently achieved by self-normalization, developed in a change-point context by Shao & Zhang (2010). Self-normalized test statistics have been applied in extreme-value contexts by Hoga (2017c) and Hoga & Wied (2017).

Table 1. $(1-\nu)$*-quantiles* c_ν *of limiting distribution in* (13) *for different levels* $\nu \in (0,1)$.

ν	0·1	0·05	0·025	0·01	0·005	0·0001
c_ν	29·6	40·1	52·2	68·6	84·6	121·9

THEOREM 1. *Suppose Assumptions* 1–3 *are met. Then, under* $\mathcal{H}_0^{(\eta)}$,

$$\mathcal{U}_n = \sup_{t\in[0,1]} \left(\frac{\left[t(1-t)\left\{\widehat{p}_n(0,t)-\widehat{p}_n(t,1)\right\}\right]^2}{\int_0^t \left[s\left\{\widehat{p}_n(0,s)-\widehat{p}_n(0,t)\right\}\right]^2 \mathrm{d}s + \int_t^1 \left[(1-s)\left\{\widehat{p}_n(s,1)-\widehat{p}_n(t,1)\right\}\right]^2 \mathrm{d}s} \right)$$

$$\longrightarrow \sup_{t\in[0,1]} \left(\frac{\left[W(t)-tW(1)\right]^2}{\int_0^t \left[W(s)-\frac{s}{t}W(t)\right]^2 \mathrm{d}s + \int_t^1 \left[W(1)-W(s)-\frac{1-s}{1-t}\left\{W(1)-W(t)\right\}\right]^2 \mathrm{d}s} \right), \quad (13)$$

in distribution as $n\to\infty$, *where* $W(\cdot)$ *denotes a standard Brownian motion.*

The null is rejected at level $\nu \in (0,1)$ if $\mathcal{U}_n$ is larger than the $(1-\nu)$-quantile of the limiting random variable in (13). Critical values are tabulated in Shao & Zhang (2010, Table 1) and are repeated in Table 1.

Theorem 1 reveals that self-normalization serves a double purpose here. It obviates the need to estimate an asymptotic variance, but also the need to estimate or even know η, because the factor $(n/k_n)^\eta$ in (11) cancels in the numerator and denominator of $\mathcal{U}_n$. This is crucial in both settings where our test can be used. In the case of serial extremal dependence where $Y_i = X_{i-h}$, estimation theory for η relies on heuristics (Ledford & Tawn, 2003, Sec. 5). When Y_i is not a lagged X_i, existing work heavily uses serially independent $(X_i, Y_i)^{\mathrm{T}}$ (Peng, 1999; Draisma et al., 2004), which is often not credible in applications.

One-time break alternatives are picked up by the numerator of $\mathcal{U}_n$ that compares subsample pseudo-estimates of $p_{n,i}^{(\eta)}$. The denominator, a functional of squared differences of the $\widehat{p}_n(s,t)$'s, obviates the need to estimate the long-run variance, and it accounts for the one-break alternative $\mathcal{H}_1^{(\eta)}$, thus avoiding the problem of nonmonotonic power caused by large estimates $\widehat{\sigma}$ that counteract large values of $G_n(t)$ under the alternative (Shao & Zhang, 2010). To see how self-normalization circumvents this problem, note that for $t = t^*$ the two summands in the denominator of $\mathcal{U}_n$ are $\mathcal{O}_{\mathrm{pr}}(1)$, because, heuristically, both integrands are based on samples without structural breaks and hence are $\mathcal{O}_{\mathrm{pr}}(1)$ uniformly in s.

Self-normalization has additional advantages: unlike for kernel-variance estimators, no tuning parameters need to be chosen; it is easy to implement because $\widehat{p}_n(0,t)$ and $\widehat{p}_n(t,1)$ need be calculated anyway; it offers good finite-sample performance (Shao & Zhang, 2010; Hoga, 2017b; Hoga, 2017c; Hoga & Wied, 2017); the computational burden is low; and in change-point contexts, no additional assumptions are needed that may otherwise be required for, e.g., valid bootstrap procedures.

The behaviour of our test under the local alternative in $\mathcal{H}_1^{(\eta)}$ is given in

THEOREM 2. *Suppose Assumptions* 1–3 *are met. Then, under* $\mathcal{H}_1^{(\eta)}$,

$$\lim_{|M|\to\infty} \lim_{n\to\infty} \mathrm{pr}\left(\mathcal{U}_n > c_\nu\right) = 1,$$

where c_ν *is the critical value for level* $\nu \in (0,1)$.

Theorem 2 shows that power under local alternatives can be arbitrarily large as the magnitude $|M|$ in $\mathcal{H}_1^{(\eta)}$ increases, so our test has non-trivial power in a $c_n^{-1/2}$-neighbourhood. This neighbourhood is the smaller and thus local power higher, the larger the value of η. This is to be expected because for higher values of η more observations lie in the joint tail, thus aiding the detection of changes. The proof of Theorem 2 reveals that the limiting distribution of $\mathcal{U}_n$ diverges at rate M^2.

Estimates of tail dependence between international stock indices obtained by Poon et al. (2003, 2004) strongly suggest the possibility of breaks in the parameter η. The following theorem shows that a test of constancy of η based on $\mathcal{U}_n$ is also consistent.

THEOREM 3. *Suppose that* $\{V_{n,i}^{\text{pre}}\}$ *and* $\{V_{n,i}^{\text{post}}\}$ *satisfy Assumptions* 1–3 *with* σ, η *replaced by* σ_{pre}, $\eta_{\text{pre}} \in (0,1]$ *and* σ_{post}, $\eta_{\text{post}} \in (0,1]$, *respectively. Also suppose that* $\mathcal{H}_0^{(\eta_{\text{pre}})}$ *and* $\mathcal{H}_0^{(\eta_{\text{post}})}$ *hold for* $\{V_{n,i}^{\text{pre}}\}$ *and* $\{V_{n,i}^{\text{post}}\}$ *with* $p_{n,1}^{(\eta_{\text{pre}})} \longrightarrow c_0^{\text{pre}} > 0$ *and* $p_{n,1}^{(\eta_{\text{post}})} \longrightarrow c_0^{\text{post}} > 0$ *as* $n \to \infty$, *respectively. Finally, suppose that*

$$V_{n,i} = \begin{cases} V_{n,i}^{\text{pre}}, & i = 1, \ldots, \lfloor nt^* \rfloor, \\ V_{n,i}^{\text{post}}, & i = \lfloor nt^* \rfloor + 1, \ldots, n, \end{cases} \qquad t^* \in (0,1),$$

satisfies Assumption 2 *with* $\eta = \min(\eta_{\text{pre}}, \eta_{\text{post}})$. *Then, under* $\mathcal{H}_A : \eta_{\text{pre}} \neq \eta_{\text{post}}$, $\lim_{n\to\infty} \text{pr}\,(\mathcal{U}_n > c_\nu) = 1$, *where* c_ν *is the critical value for level* $\nu \in (0,1)$.

Theorems 2 and 3 together show that there are two possible implications if the test statistic $\mathcal{U}_n$ falls in the critical region. In Theorem 2 the coefficient η is assumed to be constant, so that any differences in tail dependence arise only through the different asymptotic behaviour of $p_{n,i}^{(\eta)}$. Under the stronger alternative $\mathcal{H}_A$ considered in Theorem 3, however, the change in tail dependence is more pronounced as the probability $\text{pr}(X_i > b_{x,n}, Y_i > b_{y,n})$ requires a different scaling before and after the break. Thus, to be able to interpret a rejection of our test in empirical applications, one needs to check whether or not a constant η is plausible. If η appears unchanged, the break can then be pinpointed to $p_{n,i}^{(\eta)}$. An analysis of the constancy of η could follow Poon et al. (2003, 2004); see also the application in § 4.

Remark 4. So far we have considered only pairwise dependencies, but one may be interested in the joint extremal dependence of D-variate random vectors $(X_{1,i}, \ldots, X_{D,i})^{\text{T}}$, $D \geq 2$, and consider

$$p_{n,i}^{(D)} = (n/k_n)^{1/\eta_D}\, \text{pr}\left(X_{1,i} > b_{1,n}, \ldots, X_{D,i} > b_{D,n}\right), \tag{14}$$

where $b_{d,n}$ denotes the $(1 - k_n/n)$-quantile of $X_{d,i}$ $(d = 1, \ldots, D)$. Imposing Assumptions 1 and 2 for the re-defined $V_{n,i} = (X_{1,i}, \ldots, X_{D,i})^{\text{T}}$ and replacing $I_{\left(X_i > b_{x,n}, Y_i > b_{y,n}\right)}$ with $I_{\left(X_{1,i} > b_{1,n}, \ldots, X_{D,i} > b_{D,n}\right)}$ in Assumption 3, one may construct change-point tests by re-defining $\mathcal{U}_n$ in terms of

$$\widehat{p}_n^{(D)}(s,t) = \left(\frac{n}{k_n}\right)^{1/\eta_D} \frac{1}{n(t-s)} \sum_{i=\lfloor ns \rfloor + 1}^{\lfloor nt \rfloor} I_{\left(X_{1,i} > X_{1,(k_n+1)}, \ldots, X_{D,i} > X_{D,(k_n+1)}\right)}, \qquad \eta_D \in (0,1],$$

where $X_{d,(k_n+1)}$ denotes the (k_n+1)-largest value of $X_{d,1}, \ldots, X_{d,n}$ $(d = 1, \ldots, D)$. For the multivariate extension of Assumption 1, see Ramos & Ledford (2009). Charpentier & Segers (2009) summarize dependence structures in Archimedean copulas with existing limits in (14). For some important sub-families, such as the Clayton, Frank or Ali-Mikhail-Haq families, $\eta_D = \eta = 1$ is independent of the dimensionality.

2·3. *Extension to multiple breaks*

The change-point test based on $\mathcal{U}_n$ is designed to be consistent under a one-break alternative, but there may be multiple breaks in the extremal dependence structure. Shao & Zhang (2010, Sec. 2.3) outline adaptations of $\mathcal{U}_n$ that take into account the possible presence of m breaks. However, Zhang & Lavitas (2018+) show in simulations that mis-specifying the typically unknown number of possible change points m leads to a loss in power of the test of Shao & Zhang (2010), which also becomes computationally prohibitive if $m > 2$. To address these problems, Zhang & Lavitas (2018+) propose unsupervised change-point tests that do not require specifying m and have a constant computational cost. The price to pay is lower power if the number of breaks is correctly specified in Shao & Zhang's (2010) test.

We follow Zhang & Lavitas (2018+) in constructing tests against the m-break alternative

$$\mathcal{H}_{1,m}^{(\eta)} \;:\; \quad p_{n,i}^{(\eta)} = p_{n,1}^{(\eta)} + o(c_n^{-1/2}) + M_j c_n^{-1/2}, \qquad M_0 = 0,\; M_j \neq 0\; (j > 0),$$

uniformly in $i = \lfloor nt_j^* \rfloor + 1, \ldots, \lfloor nt_{j+1}^* \rfloor$, $j = 0, \ldots, m$, where $t_0^* = 0$, $t_{m+1}^* = 1$ and $\min_{j=0,\ldots,m}(t_j^* - t_{j-1}^*) > \varepsilon$ with t_j^* $(j = 1, \ldots, m)$ denoting the breakpoints. To do so, set $\widehat{S}_n(t) = c_n^{1/2} t \widehat{p}_n(0, t)$ and put

$$D(W, t_1, t_2, t_3) = \frac{1}{(t_3 - t_1)^{1/2}} \left[W(t_2) - W(t_1) - \frac{t_2 - t_1}{t_3 - t_1} \left\{ W(t_3) - W(t_1) \right\} \right],$$

$$\Xi(W, t_1, t_2, t_3) = \frac{1}{(t_3 - t_1)^2} \left(\int_{t_1}^{t_2} \left[W(s) - W(t_1) - \frac{s - t_1}{t_2 - t_1} \left\{ W(t_2) - W(t_1) \right\} \right]^2 \mathrm{d}s + \int_{t_2}^{t_3} \left[W(t_3) - W(s) - \frac{t_3 - s}{t_3 - t_2} \left\{ W(t_3) - W(t_2) \right\} \right]^2 \mathrm{d}s \right).$$

Then, our test statistic is defined as

$$\mathcal{T}_n = \mathcal{T}(\widehat{S}_n) = \sup_{\substack{\varepsilon \le r_1 < r_2 \le 1-\varepsilon \\ r_2 - r_1 \ge \varepsilon}} \frac{D^2(\widehat{S}_n, 0, r_1, r_2)}{\Xi(\widehat{S}_n, 0, r_1, r_2)} + \sup_{\substack{\varepsilon \le r_1 < r_2 \le 1-\varepsilon \\ r_2 - r_1 \ge \varepsilon}} \frac{D^2(\widehat{S}_n, r_1, r_2, 1)}{\Xi(\widehat{S}_n, r_1, r_2, 1)}. \tag{15}$$

The test statistic $\mathcal{T}_n$ can be given a more revealing form by noting that

$$D(\widehat{S}_n, 0, r_1, r_2) = \left(\frac{c_n}{r_2} \right)^{1/2} r_1 \frac{r_2 - r_1}{r_2} \left\{ \widehat{p}_n(0, r_1) - \widehat{p}_n(r_1, r_2) \right\},$$

$$\Xi(\widehat{S}_n, 0, r_1, r_2) = \frac{c_n}{r_2^2} \left(\int_0^{r_1} \left[s \frac{r_1 - s}{r_1} \left\{ \widehat{p}_n(0, s) - \widehat{p}_n(s, r_1) \right\} \right]^2 \mathrm{d}s + \int_{r_1}^{r_2} \left[(r_2 - s) \frac{s - r_1}{r_2 - r_1} \left\{ \widehat{p}_n(r_1, s) - \widehat{p}_n(s, r_2) \right\} \right]^2 \mathrm{d}s \right).$$

Similar expressions can be given for the second term on the right-hand side of (15). Under the alternative, $D(\widehat{S}_n, 0, t_1^*, t_2^*)$ picks up the changes via the two subsample estimates. At the same time $\Xi(\widehat{S}_n, 0, t_1^*, t_2^*)$ does not diverge, as the two integrals only contain estimates based on subsamples without structural breaks. This logic applies no matter how many additional breakpoints $t_3^*, \ldots, t_m^*$ there may be.

THEOREM 4. *Suppose Assumptions* 1–3 *are met. Then, (i) under* $\mathcal{H}_0^{(\eta)}$,

$$\mathcal{T}_n \longrightarrow \sup_{\substack{\varepsilon \le r_1 < r_2 \le 1-\varepsilon \\ r_2 - r_1 \ge \varepsilon}} \frac{D^2(W, 0, r_1, r_2)}{\Xi(W, 0, r_1, r_2)} + \sup_{\substack{\varepsilon \le r_1 < r_2 \le 1-\varepsilon \\ r_2 - r_1 \ge \varepsilon}} \frac{D^2(W, r_1, r_2, 1)}{\Xi(W, r_1, r_2, 1)} = \mathcal{T}(W),$$

in distribution as $n \to \infty$, *where* $W(\cdot)$ *denotes a standard Brownian motion, and (ii) under* $\mathcal{H}_{1,m}^{(\eta)}$,

$$\lim_{\min_{j=1,\ldots,m} |M_j| \to \infty} \lim_{n \to \infty} \mathrm{pr} \left\{ \mathcal{T}_n > c_{\nu, \mathcal{T}} \right\} = 1,$$

where $c_{\nu,\mathcal{T}}$ *is the* $(1 - \nu)$*-quantile of* $\mathcal{T}(W)$ *for* $\nu \in (0, 1)$.

3. SIMULATIONS

We investigate size and power of our test based on $\mathcal{U}_n$ in finite samples. In line with results in Zhang & Lavitas (2018+), unreported simulations indicate that the test based on $\mathcal{T}_n$ has comparable size, yet reduced power in case there is only one break, as we shall assume below. In the simulations for $\mathcal{T}_n$ we followed Zhang & Lavitas (2018+) and others in choosing $\varepsilon = 0{\cdot}1$. We only consider alternatives $\mathcal{H}_1^{(\eta)}$, since the power will only be higher under the more drastic tail dependence change under $\mathcal{H}_A$. All simulations are run using R version 3.4.1 (R Core Team, 2017). We use 10,000 replications throughout.

As always in applications of extreme value theory, the choice of k_n in finite samples is a delicate issue, because k_n is only specified asymptotically. We adapt the version of the plateau-finding algorithm of

Table 2. *Rejection frequencies in % for nominal levels 10%, 5% and 1% of tests based on $\mathcal{U}_n$ for trajectories of length n from model* (16)

ϕ	η	Hyp.	t^*	$n=500$			$n=2000$		
				10%	5%	1%	10%	5%	1%
0	1	$\mathcal{H}_0^{(\eta)}$		9·7	4·8	1·1	10	5·3	1·0
		$\mathcal{H}_1^{(\eta)}$	0·25	43	30	12	82	72	45
			0·50	74	62	37	99	96	86
			0·75	58	44	22	92	87	68
	1/2	$\mathcal{H}_0^{(\eta)}$		11	5·8	1·4	10	5·9	1·6
		$\mathcal{H}_1^{(\eta)}$	0·25	30	19	6·2	48	35	16
			0·50	55	43	21	73	63	42
			0·75	42	30	13	60	49	28
1/3	1	$\mathcal{H}_0^{(\eta)}$		11	5·8	1·3	10	5·6	1·2
		$\mathcal{H}_1^{(\eta)}$	0·25	32	20	6·3	69	56	29
			0·50	60	47	23	95	90	73
			0·75	47	35	16	86	77	54
	1/2	$\mathcal{H}_0^{(\eta)}$		11	6·0	1·5	11	5·8	1·3
		$\mathcal{H}_1^{(\eta)}$	0·25	24	15	4·1	45	32	13
			0·50	47	34	15	74	64	40
			0·75	37	25	10	61	50	28

Frahm et al. (2005, Sec. 4.4) used by Bücher et al. (2015), where k_n necessarily lies between roughly $n^{1/2}/2$ and n. While a choice of k_n in the neighbourhood of n is not allowed by theory, even values of around $n^{1/2}/2$ may be considered large in typical applications of extreme value theory. Hence, we propose the following adaptation.

Set $b_n = \lfloor n^{0\cdot 9}/100 \rfloor$ and choose the plateau length $m_n = (n-2b_n)^{1/2}$. Let $k_n^{\max} = \lfloor n^{0\cdot 8} \rfloor$ and $k_n^{\min} = \lfloor 10 \log n \rfloor$ represent the maximal and minimal value of k_n that one is willing to use. In a first step, set $\eta = 1$ without loss of generality. Then full-sample estimates $\widehat{p}_{n,k}^{(\eta=1)} = \widehat{p}_{n,k}^{(\eta=1)}(0,1)$ $(k = 1, \ldots, k_n^{\max} + m_n - 1 + 2b_n)$ from (11) based on different numbers of upper order statistics are smoothed by taking rolling window means of $(2b_n + 1)$ successive values. This leads to $\overline{p}_1, \ldots, \overline{p}_{k_{\max}+m_n-1}$. Then, define a plateau $p(k) = (\overline{p}_k, \ldots, \overline{p}_{k+m_n-1})$ for $k = 1, \ldots, k_n^{\max}$, and the sum of the absolute deviations between the first entry and all others by $\mathrm{SAD}(k) = \sum_{i=k+1}^{k+m_n-1} |\overline{p}_i - \overline{p}_k|$. Finally, $k_n = k_n^*$ is chosen as

$$k_n^* = \mathop{\arg\min}_{k=k_n^{\min},\ldots,k_n^{\max}} \mathrm{SAD}(k).$$

The choice of k_n^* is not affected by the particular η, since the pre-factor $(n/k_n)^\eta$ in (11) does not affect the minimum of $\mathrm{SAD}(k)$, so we could assume $\eta = 1$ without loss of generality. Setting $k_n^{\min} = 1$ is possible, but if joint exceedances above a high threshold are rare, it may happen that $\overline{p}_1 = \cdots = \overline{p}_{m_n} = 0$. Hence, $k_n^* = 1$ is chosen not because $\widehat{p}_i$-estimates do not vary much for different k_n's close to k_n^*, but only because joint high threshold exceedances did not occur.

By definition of $k_n^{\min}$ and $k_n^{\max}$, k_n^* has a rate between $\log n$ and $n^{0\cdot 8}$, which seems to agree more with typical growth rates of k_n considered in extreme value theory (Reiss & Thomas, 2007). If $k_n \sim an^b$, $a > 0$, then theory requires $b > 1 - \eta$ by Remark 2. In our simulations where η is known we could take this restriction into account, but η is unknown in practice, so it will be interesting to see how our η-independent choice k_n^* performs.

To introduce different degrees of dependence in our simulated models, consider the Joe copula

$$C(u, v) = \psi\{\psi^{-1}(u) + \psi^{-1}(v)\}, \qquad 0 \leq u, v \leq 1,$$

with generator $\psi(t) = 1 - \{1 - \exp(-t)\}^{1/\theta}$, $\theta \in [1, \infty)$. Bivariate random vectors $(X_i, Y_i)^{\mathrm{T}}$ with Joe copula exhibit asymptotic dependence with $\eta = 1$ and $\lambda = \lim_{n\to\infty} p_{n,i}^{(1)} = 2 - 2^{1/\theta}$, while the negated $(-X_i, -Y_i)^{\mathrm{T}}$ exhibit asymptotic independence with $\eta = 1/2$ and $\lim_{n\to\infty} p_{n,i}^{(1/2)} = \theta$ (Heffernan, 2000, Table 1).

We first simulate from the simple vector autoregressive model of order 1

$$(X_i, Y_i)^{\mathrm{T}} = \phi(X_{i-1}, Y_{i-1})^{\mathrm{T}} + (\epsilon_i^x, \epsilon_i^y)^{\mathrm{T}}, \tag{16}$$

where $\phi \in \{0, 1/3\}$ and $(\epsilon_i^x, \epsilon_i^y)^{\mathrm{T}}$ are independent and identically distributed. The innovations $(\epsilon_i^x, \epsilon_i^y)^{\mathrm{T}}$ are marginally standard normal. We consider two models for their dependence structure. To investigate the asymptotically dependent case $\eta = 1$, we first use the Joe copula with parameters $\theta = \theta_{n,i}$ chosen such that $\lim_{n\to\infty} p_{n,i}^{(1)} = 1/4$ under the null and $\lim_{n\to\infty} p_{n,i}^{(1)} = 1/4 + 1/2 I_{(i > \lfloor nt^* \rfloor)}$ under the alternative. Second, for the asymptotically independent case $\eta = 1/2$, we assume $(-\epsilon_i^x, -\epsilon_i^y)^{\mathrm{T}}$ to have a Joe copula with parameters $\theta = \theta_{n,i}$ chosen such that for $(\epsilon_i^x, \epsilon_i^y)^{\mathrm{T}}$, we have $\lim_{n\to\infty} p_{n,i}^{(1/2)} = 4/3$ under the null and $\lim_{n\to\infty} p_{n,i}^{(1/2)} = 4/3 + 8/3 I_{(i > \lfloor nt^* \rfloor)}$ under the alternative. So in both cases of asymptotic dependence and independence there is a threefold increase in $\lim_{n\to\infty} p_{n,i}^{(\eta)}$.

It is easy to verify Assumptions 1–3 for $(X_i, Y_i)^{\mathrm{T}}$ in (16) when $\phi = 0$. Assumption 1 is easy, yet tedious, to establish (Heffernan, 2000). Since the observations are independent and thus geometrically β-mixing, Assumption 2 is satisfied by Remark 2. Finally, from Proposition 1 and results in Heffernan (2000), Assumption 3 is met.

Table 2 shows size and power of the $\mathcal{U}_n$-test for $\phi \in \{0, 1/3\}$ and $\eta \in \{1/2, 1\}$. Size is satisfactory across all models, even for $n = 500$. As expected, power increases in sample size and proximity of the break to the middle of the time series. Also, power increases the closer ϕ is to zero. This may be explained by the confounding effects of serial dependence induced by a non-zero value of ϕ. Finally, the threefold increase in $\lim_{n\to\infty} p_{n,i}^{(\eta)}$ is more easily picked up when $\eta = 1$. This is also as expected, because when η is larger, more observations in the joint tail can be expected and thus, potentially, more evidence against the null can be provided. Theorem 2 demonstrates this theoretically. Overall, our simulations demonstrate good size of our test even for $n = 500$ and very satisfactory power for $n = 2000$. Consequently, the method of choosing k_n seems reasonable for different values of η.

4. A FINANCIAL CRISIS EXAMPLE

We examine whether extremal dependence between S&P 500 and the DAX log-returns changed during the financial crisis of 2007–8. The question is of obvious relevance to investors: spreading assets across regions is beneficial to portfolios because of its diversifying effect and, as the saying has it, diversification is the only free lunch in finance. If, however, asset prices start to co-move across asset classes or regions, then a rebalancing of the portfolio may be required. This is called diversification meltdown (Campbell et al., 2008). The recent financial crisis played out over several years, so strong prior beliefs on the location of a possible change point are hard to come by. Hence, a test for a change with an unknown location, as developed in § 2, is required here. We analyse both upper and lower tail dependence separately and come to similar conclusions.

We use log-returns calculated from adjusted daily closing prices from 2004 to 2011 of both indices, which are taken from *finance.yahoo.com*. We only keep those where data for both are available, leaving us with a total of 1991 observations $(X_1, Y_1)^{\mathrm{T}}, \ldots, (X_{1991}, Y_{1991})^{\mathrm{T}}$. Figure 1 displays the two time series. The scatter plots of the S&P 500 and DAX log-returns before and after the U.S. investment bank Lehman Brothers filed for bankruptcy protection on 15 September 2008 are shown in panels (c) and (d). The date is chosen somewhat arbitrarily, yet it was the largest bankruptcy by assets during the recent financial crisis and marked the beginning of the wildest swings in returns; see panels (a) and (b). Panel (c) clearly shows mild clustering of extremes in the upper-right and lower-left quadrants. This is more pronounced in panel (d), where both indices experience their largest and second largest returns on the same day. This gives a first indication of a possible break in the extremal dependence structure.

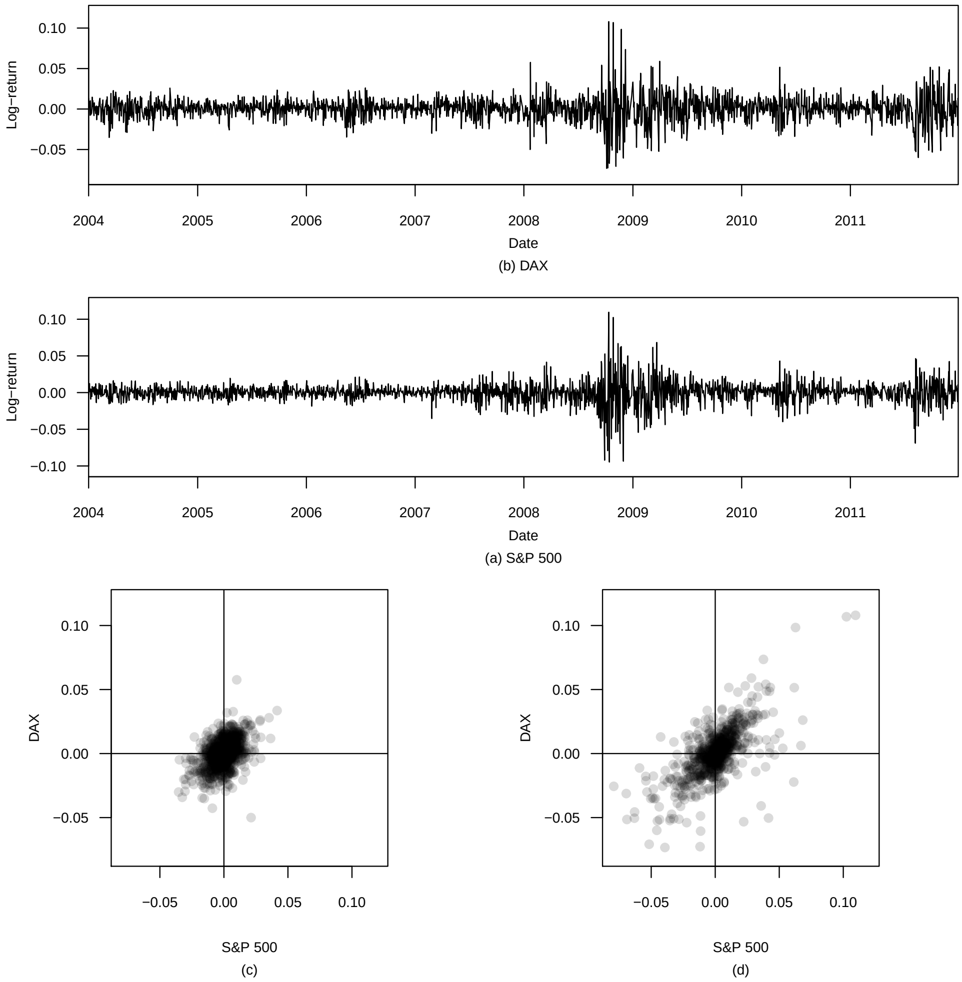

Fig. 1. Plot of log-returns of S&P 500 in (a) and DAX in (b). Scatter plots of DAX and S&P 500 log-returns before and after 15 September 2008 in (c) and (d).

Comparing panels (c) and (d) also highlights the importance of allowing for dependent data in our test. At first sight, it may appear that the marginal distributions of both the DAX and the S&P 500 returns have changed after the Lehman crisis, due to an increase of large observations in both series. However, this may just be due to dependence in volatility, that is still consistent with stationary marginal distributions, as required by our test.

We check that our test may reasonably be applied. Conveniently, we do not have to discriminate between asymptotic dependence/independence for this. We fit autoregressive moving average models with generalized autoregressive conditionally heteroscedastic errors of different orders to both S&P 500 and DAX log-returns. For each series, we determine the most suitable order using the Akaike information criterion. The parameter estimates of the best fitting models indicate stationary returns. Moreoever, Ljung–Box tests on the raw and squared standardized residuals reveal that no autocorrelation is left over, suggesting that both series are well-described by these stationary time series models, such that time-invariant marginal distributions appear plausible. As a check on this result we test for changes in the variance of the two time series. Since there appears to be little variation in the mean, we simply test for changes in the mean of the squared log-returns using the self-normalized test in Theorem 3.1 of Shao & Zhang (2010). Again, there

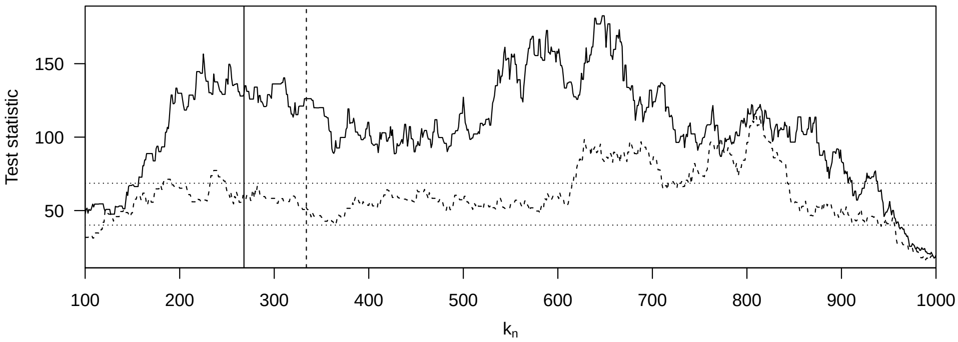

Fig. 2. Plot of $\mathcal{U}_n$ for log-returns (solid) and log-losses (dashed). 5%- and 1%-critical values (dotted horizontal lines). Vertical lines indicate choices of k_n^* for log-returns (solid) and log-losses (dashed).

is no evidence for a change in the marginal distribution, as the p-values are well above $0{\cdot}1$. Thus, we can be reasonably confident that there is no change in the marginal distribution, as required by our test.

Figure 2 plots the values of the test statistic $\mathcal{U}_n$ for different values of k_n for the log-returns and for the positive log-losses of both series. The test statistic for the null of constant lower tail dependence is above the 5%-critical value for all reasonable values of k_n and in particular for $k_n^* = 334$, which is the choice of the algorithm from § 3. The test statistic for constant upper tail dependence yields even more significant results, with the test statistic well above the 1%-critical value, also for $k_n^* = 268$. This is convincing evidence for a break in the lower and in the upper tail dependence of the bivariate time series.

Four questions remain: first, in which direction was the break, i.e., are extremal co-movements more or less likely after it? Second, when did it occur? Third, were there possibly more breaks? Fourth, can the change in tail dependence be attributed to differences merely in $p_{n,i}^{(\eta)}$, or has there been a change in η?

Figure 3 plots $G_n(t)$ from (12) with $\eta = 1$ and provides some answers to the first three questions. The particular value of η has no consequence for the following conclusions, since it just leads to a different scaling of $G_n(t)$. For both tests, $G_n(t)$ is exclusively negative, so extremal dependence in the upper and the lower tail of S&P 500 and DAX log-returns has likely intensified during the crisis, which is evidence for diversification meltdown. The minimum of $t \mapsto G_n(t)$ is attained on 16 July 2008 for the solid line and 9 July 2007 for the dashed line. These dates indicate a break in the extremal dependence even before the Lehman bankruptcy on 15 September 2008. Splitting the bivariate sample at both minima and testing for another break in the upper and lower extremal dependence in the respective subsamples, we find no evidence for additional breaks. Some further evidence for the presence of only one break comes from applying the $\mathcal{T}_n$-test of § 2·3. Unreported simulations show that this test has lower power than the $\mathcal{U}_n$-test under the one-break alternative. Indeed, applying the $\mathcal{T}_n$-test to both the bivariate log-losses and log-returns, with $\varepsilon = 0{\cdot}1$ and k_n^*'s as above, we only find evidence for a break in the upper tail dependence at significance level 10%; the p-value of a test of constant lower tail dependence is above $0{\cdot}1$.

As for the fourth question, panels (c) and (d) in Fig. 1 suggest that before the crisis the returns may have been tail independent with $\eta < 1$, only to become tail dependent with $\eta = 1$ afterwards. We split the sample at the minimum of $G_n(t)$, i.e., 16 July 2008 for the upper joint tail and 9 July 2007 for the lower joint tail. We use estimation and inference methods for η that, to the best of our knowledge, only exist for serially independent data. Thus, the following conclusions must be interpreted with care. Estimates of η based on the estimator in Coles et al. (1999, Sec. 3.3.2) suggest values of η of $0{\cdot}71$ and $0{\cdot}66$ for the pre-break lower and upper tails, and values of $0{\cdot}82$ and $0{\cdot}81$ for the post-break period. The respective confidence intervals for η in the pre-/post-break periods do not include the post-/pre-break estimates. This suggests a break in η. This conclusion is supported by the test of $\eta = 1$ in Reiss & Thomas (2007,

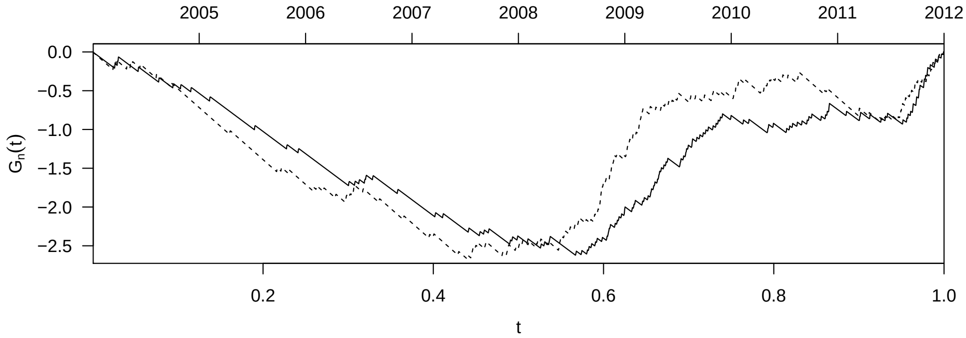

Fig. 3. Plot of $t \mapsto G_n(t)$ for log-returns (solid; $k_n^* = 268$) and log-losses (dashed; $k_n^* = 334$).

Sec. 13.3). For both the upper and lower joint tail, the test rejects the null hypothesis of tail dependence in the pre-crisis sample with p-values below $0{\cdot}05$ for a range of possible thresholds. For the post-crisis sample however, tail dependence can no longer be rejected. For both break dates the number of observations either side of the break is roughly equal, so that the non-rejection of $\eta = 1$ after the breaks is not due to a mere lack of data.

Our results suggest one break in the coefficient of tail dependence given tail independence η somewhere around the beginning of the crisis, after which extremal co-movements in the upper and lower joint tail of both indices have become more frequent.

Acknowledgement

The author is indebted to an Associate Editor, two referees and Christoph Hanck for numerous suggestions that improved the paper. Thanks also go to Dominik Wied for sharing his R code. Support of Deutsche Forschungsgemeinschaft is gratefully acknowledged.

Supplementary material

Supplementary material available at *Biometrika* online includes proofs of the theoretical results and simulations comparing our test with the test for correlation changes of Wied et al. (2012).

References

Aue, A., Hörmann, S., Horváth, L. & Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* **37**, 4046–4087.

Bae, K. H., Karolyi, G. A. & Stulz, R. M. (2003). A new approach to measuring financial contagion. *Rev. Financ. Stud.* **16**, 717–763.

Bücher, A., Jäschke, S. & Wied, D. (2015). Nonparametric tests for constant tail dependence with an application to energy and finance. *J. Economet.* **187**, 154–168.

Campbell, R. A. J., Forbes, C. S., Koedijk, K. G. & Kofman, P. (2008). Increasing correlations or just fat tails? *J. Empirical Finance* **15**, 287–309.

Charpentier, A. & Segers, J. (2009). Tails of multivariate Archimedean copulas. *J. Mult. Anal.* **100**, 1521–1537.

Coles, S., Heffernan, J. & Tawn, J. A. (1999). Dependence measures for extreme value analyses. *Extremes* **2**, 339–365.

Davis, R. A. & Mikosch, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli* **15**, 977–1009.

Davis, R. A., Mikosch, T. & Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. *J. Economet.* **170**, 142–152.

DAVIS, R. A., MIKOSCH, T. & ZHAO, Y. (2013). Measures of serial extremal dependence and their estimation. *Stoch. Proces. Appl.* **123**, 2575–2602.

DRAISMA, G., DREES, H., FERREIRA, A. & DE HAAN, L. (2004). Bivariate tail estimation: Dependence in asymptotic independence. *Bernoulli* **10**, 251–280.

DREES, H. & ROOTZÉN, H. (2010). Limit theorems for empirical processes of cluster functionals. *Ann. Statist.* **38**, 2145–2186.

EMBRECHTS, P., MCNEIL, A. & STRAUMANN, D. (2002). Correlation and dependence in risk management: Properties and pitfalls. In *Risk Management: Value at Risk and Beyond*, M. Dempsey, ed. Cambridge: Cambridge University Press, pp. 176–223.

FASEN, V., KLÜPPELBERG, C. & SCHLATHER, M. (2010). High-level dependence in time series models. *Extremes* **13**, 1–33.

FRAHM, G., JUNKER, M. & SCHMIDT, R. (2005). Estimating the tail-dependence coefficient: Properties and pitfalls. *Insurance Math. Econom.* **37**, 80–100.

HEFFERNAN, J. E. (2000). A directory of coefficients of tail dependence. *Extremes* **3**, 279–290.

HOGA, Y. (2017a). Change point tests for the tail index of β-mixing random variables. *Economet. Theory* **33**, 915–954.

HOGA, Y. (2017b). Monitoring multivariate time series. *J. Mult. Anal.* **155**, 105–121.

HOGA, Y. (2017c). Testing for changes in (extreme) VaR. *Economet. J.* **20**, 23–51.

HOGA, Y. & WIED, D. (2017). Sequential monitoring of the tail behavior of dependent data. *J. Statist. Plann. Inference* **182**, 29–49.

LEDFORD, A. W. & TAWN, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169–187.

LEDFORD, A. W. & TAWN, J. A. (1997). Modelling dependence within joint tail regions. *J. R. Statist. Soc. B* **59**, 475–499.

LEDFORD, A. W. & TAWN, J. A. (2003). Diagnostics for dependence within time series extremes. *J. R. Statist. Soc. B* **65**, 521–543.

LIEBSCHER, E. (2005). Towards a unified approach for proving geometric ergodicity and mixing properties of nonlinear autoregressive processes. *J. Time Ser. Anal.* **26**, 669–689.

MEITZ, M. & SAIKKONEN, P. (2008). Stability of nonlinear AR-GARCH models. *J. Time Ser. Anal.* **29**, 453–475.

PENG, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Statist. Probab. Lett.* **43**, 399–409.

POON, S.-H., ROCKINGER, M. & TAWN, J. A. (2003). Modeling extreme value dependence in international stock markets. *Statist. Sinica* **13**, 929–953.

POON, S.-H., ROCKINGER, M. & TAWN, J. A. (2004). Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *Rev. Financ. Stud.* **17**, 581–610.

R CORE TEAM (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

RAMOS, A. & LEDFORD, A. W. (2009). A new class of models for bivariate joint tails. *J. R. Statist. Soc. B* **71**, 219–241.

REISS, R.-D. & THOMAS, M. (2007). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Basel: Birkhäuser, 3rd ed.

SHAO, X. & ZHANG, X. (2010). Testing for change points in time series. *J. Am. Statist. Assoc.* **105**, 1228–1240.

SIBUYA, M. (1960). Bivariate extreme statistics. *Ann. Inst. Statist. Math.* **11**, 195–210.

VOGELSANG, T. J. (1997). Wald-type tests for detecting breaks in the trend function of a dynamic time series. *Economet. Theory* **13**, 818–849.

VOGELSANG, T. J. (1999). Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series. *J. Economet.* **88**, 283–299.

WIED, D., KRÄMER, W. & DEHLING, H. (2012). Testing for a change in correlation at an unknown point in time using an extended functional delta method. *Economet. Theory* **28**, 570–589.

ZHANG, T. & LAVITAS, L. (2018+). Unsupervised self-normalized change-point testing for time series. Forthcoming in *J. Am. Statist. Assoc.* (DOI: 10.1080/01621459.2016.1270214) , 1–13.