Tail event driven networks of SIFIs

Cathy Yi-Hsuan Chen, Wolfgang Karl Härdle, Yarema Okhrin

Angaben zur Veröffentlichung / Publication details:

Chen, Cathy Yi-Hsuan, Wolfgang Karl Härdle, and Yarema Okhrin. 2019. "Tail event driven networks of SIFIs." *Journal of Econometrics* 208 (1): 282–98. https://doi.org/10.1016/j.jeconom.2018.09.016.

Tail event driven networks of SIFIs

Cathy Yi-Hsuan Chen [a], Wolfgang Karl Härdle [a,b,*], Yarema Okhrin [c]

[a] *Humboldt-Universität zu Berlin, School of Business and Economics, Unter den Linden 6, 10099 Berlin, Germany*
[b] *SKBI, Singapore Management University, Singapore*
[c] *University of Augsburg, Faculty of Business and Economics, Chair of Statistics, 86159 Augsburg, Germany*

1. Introduction

Systemic risk threatens financial stability and the functioning of financial markets due to shocks on liquidity, reduced market confidence and willingness of risk taking. A growing body of literature discusses macroprudential risk management approaches to study systemic risk with two goals: ensuring financial stability and quantifying a risk charge proportional to the relative systemic contribution. Both goals are targeted to inject additional capital into the financial system to make it more resilient. However, as pointed out by Bluhm and Krahnen (2014) macroprudential monitoring is still at a very early stage, quantifying the magnitude of systemic risk and identifying systemically relevant contributors need more scientific analysis. Without such effort, the supervisory authorities face difficulties either to set proper capital requirements as a risk buffer against adverse shocks for every financial institution or to calculate additional risk charges for SIFIs for their extra negative externalities on the financial system.

In order to understand the interconnectedness among SIFIs from a system-wide perspective, one has to study the measure, the degree, and asymmetric nature of systemic risk. To do so we propose a quantitative and system-wide framework based on a topological network methodology. Network analysis is quite capable of portraying the interplay

* Corresponding author at: Humboldt-Universität zu Berlin, School of Business and Economics, Unter den Linden 6, 10099 Berlin, Germany.
E-mail addresses: cathy.chen@hu-berlin.de (C.Y.-H. Chen), haerdle@hu-berlin.de (W.K. Härdle), yarema.okhrin@wiwi.uni-augsburg.de (Y. Okhrin).

among financial institutions and measuring their interconnectedness (see Diebold and Yılmaz (2014) and Barigozzi and Brownlees (2015)). Summarizing their arguments, interconnectedness of financial institutions on the interbank market is an absolute key to understanding systemic risk. Interconnectedness captures the situations when financial distress in one institution subsequently raises the likelihood of financial distress in other institutions because of their network of contractual relations and interbank lending among them, leading to a "too-interconnected-to-fail" situation. The complexity of the connections brings challenges to researchers, a tailored network structure representing the interconnections of a large panel as a graph or an adjacency matrix though provides insight into the mechanics of financial distress.

This study proposes modern network techniques for the analysis of the network effect on the tail events in five aspects. First, we concentrate on the network of SIFIs due to their systemic relevance. The interconnectedness of SIFIs shall be quantified and kept monitored. Second, we propose and argue that a network (or "adjacency matrix") should be based on tail event related profiles. The adjacency matrix is constructed by the similarity of risk profiles of node pairs. The network being built from a risk profile comprising of various tail event variables is more capable to portray stress related structure and to represent diverse risk contents. The tail event risk profiles are conventionally allocated into contagion and diversification baskets, leading to a positive adjacency matrix A^+ and a negative one A^-. The positive (negative) network factor is therefore constructed through the estimated A^+ (A^-). A joint spacings variance ratio (SVR) test verifies the significance of similarities in an asymmetric neighborhood of zero and can provide additional evidence in favor of two adjacency matrices.

Third, the SIFIs identified by the Financial Stability Board (FSB) are allocated to different buckets depending on their scores of policy measures proposed by Basel Committee on Banking Supervision (BCBS) to admit that each SIFI potentially creates different degrees of contagion. The five buckets correspond to the additional loss absorption capacities from 1% to 3.5% to ensure the sufficiency of their common equities in case of the default. However, the methodology proposed by BCBS depends very much on the choice of policy measures and threshold of scores. Therefore, we present a parsimonious and intuitive metric for the "aggregate risk" in a network, the resulting systemic risk scores can be used to monitor systemic vulnerability. The risk decomposition of the aggregate systemic risk quantifies the risk contribution of each SIFI. For a supervisory purpose, a SIFI with a high degree of connectedness may thus be monitored more carefully for higher risk increments. In a nutshell, this analytic decomposition enables identifying the source of systemic vulnerabilities.

Fourth, the responses/outputs at the network vertices constitute an ultrahigh dimensional vector, that is, the returns of the defined SIFIs. The interest of our analysis lies in the extremes or higher moments, see Härdle et al. (2016) and Fan et al. (2016). Given an adjacency matrix one is interested in how a stress impulse on one node (or a collection of these) propagates through a network. A basic idea is seeded in the CoVaR concept of Adrian and Brunnermeier (2016) but the situation we look at requires an extension to a dynamic framework. In order to study such tail event risk transmission in a dynamic context we propose the *Tail Event driven Network Quantile Regression* (TENQR) model, extending the variation about the mean analysis of Zhu et al. (2017). The technique presented here is similar to Zhu et al. (2018), but different in the concrete construction of the adjacency matrix as we will see later. In a TENQR framework the SIFI returns given a range of quantile levels or τ values can be governed by the network factor as an indicator of the degree of connectedness in the system. Furthermore, we control for the autoregressive impact and the market-wide covariates, while the individual node-specific features are captured by employing the fixed-effects panel data framework. The TENQR model does reflect the nature of systemic risk that has (i) large impact, particular in the downside constituting an asymmetric reaction; (ii) widespread coverage that can be examined through a panel or stack model; (iii) a ripple effect being detected by an intertemporal investigation.

Fifth, we analyze the geographic variation with respect to systemic risk. For this purpose we take the panel quantile regressions for the SIFIs from the US, Europe and Asia. In doing so, one may detect which region is more sensitive to the network factor. Some regions may turn out to be risk transmitters, while others are risk recipients. Having established TENQR one contributes to a "manageable" systemic risk: the supervisors are able to rank the systemic importance for each SIFIs, to measure the resulting connectedness in a system, and to evaluate the impact of network on the conditional quantile of a response.

The empirical evidence can be summarized as follows. The negative similarities based on adverse movements of the risk profiles can be observed when pairing any SIFI in the US or EU with one in Asia. Hence, the Asian region contributes to risk diversification documented also in the risk decomposition analysis and are indeed the major active elements of the negative network factor. A joint SVR test supports the design of adjacency in this study by proving the statistical significance of the discrepancy between the binary elements. The quantile curve w.r.t. the positive network factor exhibits a downward slope (in EU and in the US), or a U-shape (in Asia) indicating a tail-event based contagion. However, for the case of the negative network factor, the monotonic increasing quantile curves document a less gain from the risk diversification in the lower tail than in the upper tail.

The paper is structured as follows. In the next section we discuss the construction of the adjacency matrix and its relevance for the systemic risk score. Section 3 contains the details for the TENQR network quantile regression. Section 4 concludes.

2. Adjacency matrix and systemic risk score

2.1. Financial characteristics of SIFIs

As institutions grow in size, interconnectedness and complexity, they profit by lower funding cost and economy of scale. However, the moral-hazard problems or so called "negative externalities" subsequently raise because governments have

Table 1
The overview of SIFIs.

Index	Name of SIFI	Firm size	Debt ratio	Bucket	Country
1	JP MORGAN CHASE	21.506	0.261	4	U.S.
2	BANK OF AMERICA	21.446	0.302	2	U.S.
3	BANK OF NEW YORK MELLON	19.499	0.095	1	U.S.
4	CITIGROUP	21.359	0.300	3	U.S.
5	GOLDMAN SACHS	20.624	0.509	2	U.S.
6	MORGAN STANLEY	20.501	0.417	2	U.S.
7	STATE STREET	19.106	0.153	1	U.S.
8	WELLS FARGO	20.980	0.183	1	U.S.
9	ROYAL BANK OF SCTL	21.588	0.252	1	U.K.
10	BARCLAYS	21.604	0.286	3	U.K.
11	HSBC	21.682	0.127	4	U.K.
12	STANDARD CHARTERED	20.136	0.187	1	U.K.
13	BNP PARIBAS	21.684	0.136	3	France
14	CREDIT AGRICOLE	21.489	0.211	1	France
15	SOCIETE GENERALE	21.184	0.139	1	France
16	DEUTSCHE BANK	21.630	0.200	3	Germany
17	UNICREDIT	20.929	0.360	1	Italy
18	ING GROEP	21.156	0.103	1	Netherlands
19	SANTANDER	21.158	0.368	1	Spain
20	NORDEA BANK	20.476	0.326	1	Sweden
21	CREDIT SUISSE GROUP N	20.744	0.339	2	Switzerland
22	UBS GROUP	21.008	0.251	1	Switzerland
23	BANK OF CHINA	21.200	0.160	1	China
24	ICBC	21.508	0.089	1	China
25	CHINA CON.BANK	21.281	0.092	1	China
26	MITSUBISHI UFJ	21.533	0.159	2	Japan
27	MIZUHO	21.247	0.233	1	Japan
28	SUMITOMO.MITSUI	21.044	0.125	1	Japan

Note: Debt ratio, a ratio of total debt to total asset, and bank size, as log value of total assets denominated in the US dollar, are shown as their mean value during sample period (2007–2015). The buckets assigned by BCBS correspond to required level of additional common equity loss absorbency as a percentage of risk-weighted assets from 3.5% (Bucket 5), 2.5% (Bucket 4), 2.0% (Bucket 3), 1.5% (Bucket 2) to 1% (Bucket 1).

been forced to use public funds to support distressed financial institutions, leading to a "too-interconnected-to-fail" or "too-big-to-fail" situation. Due to the importance of SIFIs for market stability the analysis of interconnectedness and identification of major risk contributors are important issues. Here, we analyze 28 global SIFIs (or called G-SIB) listed and updated in Nov. 2015 by FSB, but disregard the Agriculture Bank of China and Banque Populaire CE, due to their relative meager data availability. Daily data has been collected ranging from Jan. 2007 to Dec. 2015.

In Table 1, we list the names of the SIFIs with the corresponding index numbers assigned in this research, and summarize the bank-specific attributes such as debt ratio, firm size, country where the headquarter is located and the buckets assigned by the Basel Committee on Banking Supervision (BCBS). Debt ratio, a ratio of total debt to total asset, captures the fragility of a bank, while the size, as total assets, proxies for the bank being too big to fail. In particular, size risk is the most determinant standalone bank risk in relation to systemic risk (Laeven et al., 2015).

2.2. Similarity matrix

Graph theory is very useful to represent and visualize complexity of interactions. A graph is comprised of a series of nodes/vertices and the edges, referring to connect or join two nodes i, j. In this study, each node is represented as a particular SIFI, while the edge between two nodes indicates their dependence. The properties of a graph can be expressed by its adjacency matrix A, a square matrix comprising the elements of a_{ij} with value equal to one if an edge connects to node i and j and zero otherwise. In social networks such as Facebook or Twitter, a friendship can be naturally defined as it could be for example a follower–followee relationship. However, for institutional networks, defining an adjacency matrix is not as intuitive or trivial as what can be done in individual networks. One may need additional prior knowledge regarding counterparty linkages such as their contractual obligations, interbank lending and the assets they hold mutually, which are relatively hard to be gathered in time and are of very low-frequency data type. Therefore, Diebold and Yılmaz (2014) proposed to use daily stock returns for its forward-looking assessment advantage and being able to reflect the health of SIFI in time. The mapping to a network comes via variance decomposition of volatility. Concentration on volatility though falls short when quantifying the dynamics of tail event related measures in a more general non-Gaussian context.

We therefore propose to construct an adjacency matrix from risk profiles. As risk profile we take the vectors of conditional (marginal) Expected Shortfalls (CoES). This is consistent with Basel committee's recommendations and can be seen as an extension of the CoVaR concept of Adrian and Brunnermeier (2016), Härdle et al. (2016) and Hautsch et al. (2014). The CoES of the SIFI i conditional on the VaR exceedance of the SIFI j at α level is defined as:

$$\mathrm{CoES}_{ij,t}(\alpha) = \mathrm{E}[Y_{i,t} | Y_{j,t} < VaR_{j,t}(\alpha)]. \tag{1}$$

Each SIFI i is characterized at every time point t by a $N = 28$ dimensional risk profile vector $X_{i,t} = \{\text{CoES}_{ij,t}\}_{j=1,\ldots,N}$. To be specific, $X_{i,t}$ contains $N-1$ expected shortfalls conditional on the stress from $N-1$ counterparts and the expected shortfall itself while $i = j$. $\text{CoES}_{ij,t}(\alpha)$ and $VaR_{j,t}$ at 95% level are estimated respectively by the (conditional) mean and the sample quantile of the return series $Y_{j,t-h}, \ldots, Y_{j,t}$.

Unlike the CoVaR, the CoES takes into account all possible losses that exceed the severity level corresponding to the VaR of the counterpart. Besides satisfying a coherent risk measure, the CoES is continuous w.r.t α regardless of the underlying distribution. Therefore, it is not sensitive to small changes in α.

A pair of SIFIs is similar if their risk profiles share a certain degree of closeness. Take the risk profile vector comprised of CoES risk covariates and calculate the cosine similarity at each time point t:

$$\rho_{ij,t} = \frac{X_{i,t}^{\top} X_{j,t}}{\|X_{i,t}\| \|X_{j,t}\|} \text{for } j \neq i,\ i = 1, \ldots, N,\ t = 1, \ldots, T, \tag{2}$$

The risk profile similarity in (2) is analogous to the Pearson correlation coefficient, and is the dot (scalar or inner) product of the normalized and centered risk profile vectors. For any pair (i, j), as long as the normalized risk vectors move to the same direction, the cosine of the angle between them shall be very small, therefore, $\rho_{ij,t}$ will approach one (analogous to positive ΔCoVaR or positive tail dependence). Likewise, moving to reverse directions will result in a larger angle and a smaller or even negative $\rho_{ij,t}$ if the angle is higher than $\pi/2$ (analogous to negative ΔCoVaR or negative tail dependence). The appealing feature is that it covers a wide range of tail risks.

In order to understand the situations in which circumstance the similarity $\rho_{ij,t}$ will be negative, let us consider a simple case of a mean zero multivariate Normal distribution with standard normal marginals so that the covariance matrix equals to the correlation matrix $\{\varrho_{ij}\}$ with $i, j = 1, \ldots, N$. The CoES vector for SIFI i is $\mathsf{E}[Y_{i,t}|Y_{j,t} \leq \Phi^{-1}(\alpha)], \quad j = 1, \ldots, N$. In this case, for $i = j$ we obtain the well known formula, see Borak et al. (2013):

$$\mathsf{E}[Y_{i,t}|Y_{i,t} \leq \Phi^{-1}(\alpha)] = \frac{-1}{1-\alpha}\varphi\{\Phi^{-1}(\alpha)\}, \tag{3}$$

where φ and Φ are pdf and cdf of the standard normal distribution. For $i \neq j$, there is no simple analytical formula available like (3) although the risk profile is easy to estimate empirically (at least at the 5% confidence level that we use). The reason is that for the joint cdf F_{Y_i,Y_j} of two random variables (Y_i, Y_j) the expected value of Y_i conditional on Y_j lower than a certain quantity is

$$\mathsf{E}[Y_i|Y_j \leq y_j] = \int_0^{\infty} \frac{1 - F_{Y_i,Y_j}(y_i, y_j)}{F_{Y_j}(y_j)} dy_i - \int_{-\infty}^{0} \frac{F_{Y_i,Y_j}(y_i, y_j)}{F_{Y_j}(y_j)} dy_i \tag{4}$$

cannot be simplified easily even in the bivariate normal case. More insight is possible if we pick up the idea of Adrian and Brunnermeier (2016) and only look at $\mathsf{E}[Y_i|Y_j = y_j]$. Further simplification is feasible in the multivariate normal case since then the conditional distribution of Y_i given $Y_j = y_j$ will be $N(\varrho_{ij}, 1 - \varrho_{ij}^2)$. The similarity thus can be simply calculated from the risk profile vector

$$\mathsf{E}[Y_i|Y_j = \Phi^{-1}(\alpha)] = \varrho_{i,j}\Phi^{-1}(\alpha).$$

If we refer to the similarity defined in (2), we find

$$\rho_{ij} = \sum_{k=1}^{N} \varrho_{i,k}\varrho_{k,j}\{\Phi^{-1}(\alpha)\}^2$$

The illustration based on the Gaussian and Adrian and Brunnermeier (2016) framework sheds some insights in (2), that is, in what the extent the pairwise ρ_{ij} can be negative, although the CoES lacks an analytical expression for the non-Gaussian case.

For the purpose of visualization, we snapshot eight situations of the cosine similarity in (2) at every 250-day and display them in Fig. 1. The colors used to represent the degree of similarity vary from negative (blue) to positive correlation (yellow) in the grids. Through 2007–2009, one observes the similarity moving to a wide coverage of yellow color to signal an increasing (decreasing) interdependence on positive (negative) side. During the severest time of system stress, most of financial institutions are tightly tail connected except for the Asian SIFIs. As the tension for crisis has gradually mitigated, one sees a reduced connectedness in 2010. The European debt crisis in 2011–2012 boosted the similarities again, particularly for the European SIFIs (nodes 11–22). Fig. 1 shows that the similarity matrix indeed varies over time. The incorporation of this dynamic effect is left for further research. To promptly capture an approximate dynamics of risk similarity, the CoES are estimated during a given window size. A 250-day horizon $h = 250$ is employed to estimate CoES in (1).

2.3. Adjacency matrix

By its very construction the pairwise similarities (2) do not reveal equal severity: some are profound but some are not. It is not advisable to take all pairwise similarities into account if they are not beyond a certain threshold, an observation also

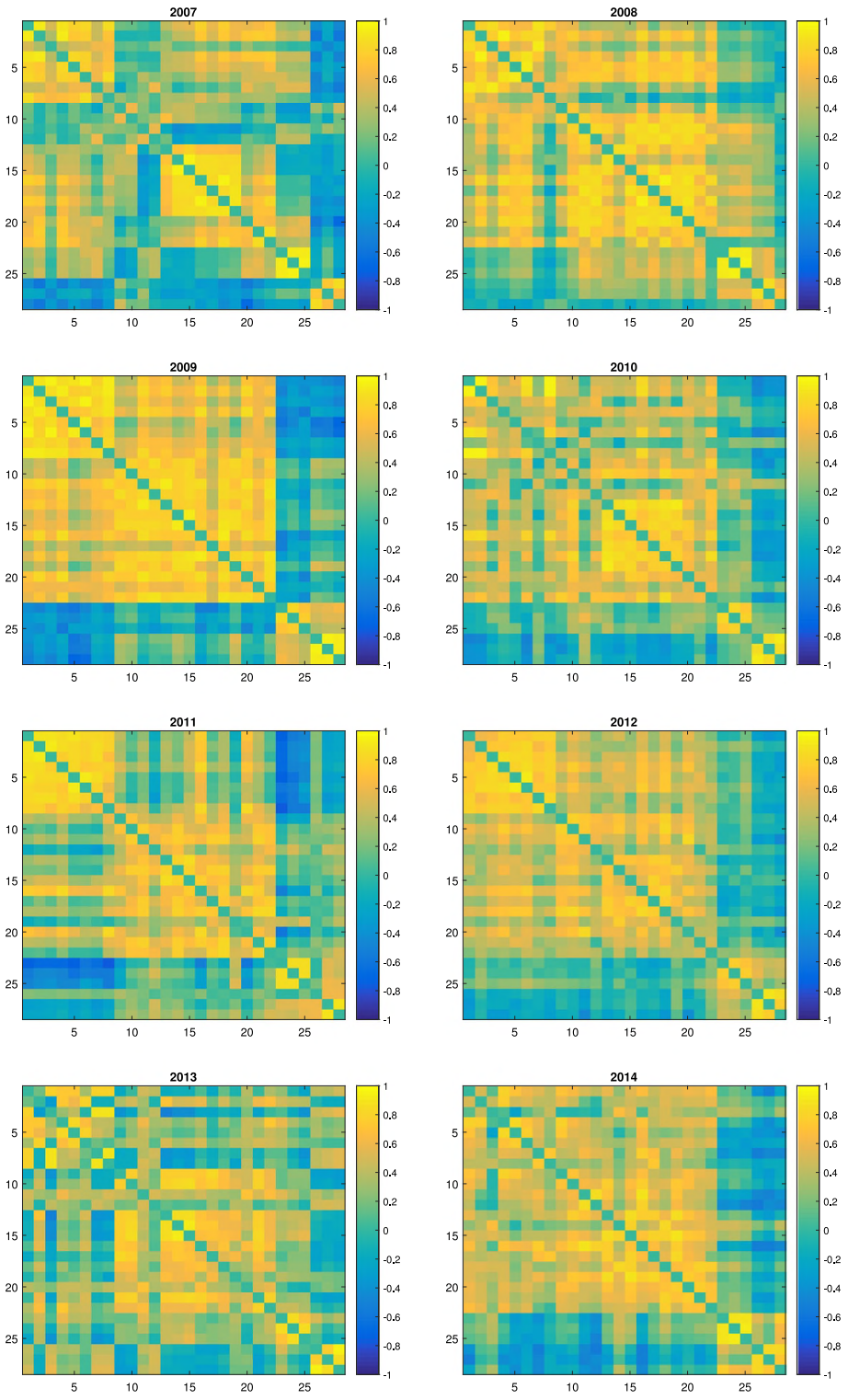

Fig. 1. Similarity matrix.

made by Härdle et al. (2016), Hautsch et al. (2014) and Barigozzi and Brownlees (2015). Therefore, to achieve manageable and interpretable structures, dimension reduction techniques are employed.

The simplest network structure is based on binary weights representing the links between the nodes, with one (zero) used to represent a link (isolation). Systemic risk, however, is induced by positive interdependencies, whereas the negative ones are benefiting from a risk diversification. Thereby, the instability indeed is caused by positive rather than negative interdependence, suggesting an asymmetric impact. The necessity to treat positive and negative correlations (the entries of the similarity matrix) differently makes therefore such a breakdown into a simple binary structure infeasible. We therefore propose an *asymmetric breakpoint approach* between positive and negative correlation to fulfill the need of creating more groups with distinguishable positive and negative linkages.

2.3.1. Asymmetric breakpoint approach

The question is how to determine the adjacency relation among the SIFIs from the similarity profile containing concurrent (positive) or disparate (negative) risk profile correlation. For this purpose, we partition the correlations into three groups: the positive, the negative and zero correlation. Therefore, a direct application of the technique of Ng (2006) is not possible. Thus we treat large negative correlations and large positive correlations asymmetrically. "Large" here means a manifest magnitude of estimated correlation. Mathematically speaking, their values must exceed the estimated threshold.

The assignment of the correlations to one of three groups is based on a breakpoint technique. Algebraically, let $\rho = (\rho_1, \rho_2, \ldots, \rho_n)^\top$ to be the vector of ordered similarities in (2) and $\rho_1 < \rho_2 < \cdots < \rho_n$ where $n = N(N-1)/2$. Extending Ng (2006), we first divide the vector of ordered similarities into the positive vector $\rho^+ = \left(\rho_1^+, \rho_2^+, \ldots, \rho_{n_1}^+\right)^\top$ and the negative vector $\rho^- = \left(\rho_1^-, \rho_2^-, \ldots, \rho_{n_2}^-\right)^\top$, and $n_1 + n_2 = n$. In our sample $n_1 > n_2$ is the common case. Fig. 6 displays the proportion n_2/n as a time-series 2008–2015. One sees that only about 20% of the SIFI pairs exhibit negative correlations.

We exploit the uniform spacings to address the number, homogeneity and magnitude of positive v.s. negative correlated pairs. Applying the cdf of standard normal distribution Φ to ordered correlations for approximate uniformity, we have

$$\phi^+ = (\phi_1^+, \phi_2^+, \ldots, \phi_{n_1}^+)^\top = (\Phi(\sqrt{N}\rho_1^+), \ldots, \Phi(\sqrt{N}\rho_{n_1}^+))^\top,$$
$$\phi^- = (\phi_1^-, \phi_2^-, \ldots, \phi_{n_2}^-)^\top = (\Phi(\sqrt{N}\rho_1^-), \ldots, \Phi(\sqrt{N}\rho_{n_2}^-))^\top,$$

Since $\rho_k^+ \in [0, 1]$ and $\rho_k^- \in [-1, 0]$, it follows that $\phi^+ \in [0.5, 1]$ and $\phi^- \in [0, 0.5]$.

The edges of the network are constructed based on large spacings between two subsequent correlations to indicate a large mean-shift in the original correlations or a slope shift in the transformed values. We therefore denote $\Delta_k^+ = \phi_k^+ - \phi_{k-1}^+$ and $\Delta_k^- = \phi_k^- - \phi_{k-1}^-$ as the positive spacings and the negative ones, respectively. The basic idea is to split the sequence of spacings into three subsets using a cluster approach. Let θ^- be the fraction of spacings which corresponds to highly negative correlations; θ^+ be the fraction of spacings which separates highly positive correlations. For the spacings in the positive group, a minimization of the total sum of squared residuals in small positive S^+ v.s. large positive L^+ results in:

$$\hat{\theta}^+ = \underset{\theta^+ \in [\underline{\theta}, \bar{\theta}]}{\operatorname{argmin}} \sum_{k=1}^{[\theta^+ n_1]} (\Delta_k^+ - \mu_S^+)^2 + \sum_{k=[\theta^+ n_1]+1}^{n_1} (\Delta_k^+ - \mu_L^+)^2 \tag{5}$$

with $\mu_S^+ = \frac{1}{[\theta^+ n_1]} \sum_{k=1}^{[\theta^+ n_1]} \Delta_k^+$ and $\mu_L^+ = \frac{1}{n_1 - [\theta^+ n_1]} \sum_{k=[\theta^+ n_1]+1}^{n_1} \Delta_k^+$. For the spacings in the negative group, optimization of the total sum of squared residuals in small negative S^- v.s. large negative L^- is:

$$\hat{\theta}^- = \underset{\theta^- \in [\underline{\theta}, \bar{\theta}]}{\operatorname{argmin}} \sum_{k=1}^{[\theta^- n_2]} (\Delta_k^- - \mu_S^-)^2 + \sum_{k=[\theta^- n_2]+1}^{n_2} (\Delta_k^- - \mu_L^-)^2 \tag{6}$$

with $\mu_S^- = \frac{1}{[\theta^- n_2]} \sum_{k=1}^{[\theta^- n_2]} \Delta_k^-$ and $\mu_L^- = \frac{1}{n_2 - [\theta^- n_2]} \sum_{k=[\theta^- n_2]+1}^{n_2} \Delta_k^-$

where Δ_k are ordered spacings and $[\theta n]$ is the integer part of θn. We have taken $\underline{\theta} = 0.1 = 1 - \bar{\theta}$ and found stable results for this choice. We obtain the optimal break fractions $\hat{\theta}^+$ resulting in a minimum total sum of variances from positive subgroups to make the correlations in a given group as homogeneous as possible. Analogously, the optimal break fraction $\hat{\theta}^-$ leads to a minimum total sum of variances from negative subgroups.

Having estimated the break fractions $\hat{\theta}^+, \hat{\theta}^-$ at each time point, we define the adjacency matrix, A, with the elements:

$$a_{ij} = \begin{cases} 1 & \text{if } \rho_{v^+(i,j)}^+ > \rho_{\hat{\theta}^+}^+ \\ -1 & \text{if } \rho_{v^-(i,j)}^- < \rho_{\hat{\theta}^-}^- \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where $v^+(i,j)$ and $v^-(i,j)$ are mappings which assign to each pair (i,j) of SIFIs the index of the ordered positive or negative correlations. $\rho_{\hat{\theta}^+}^+$ and $\rho_{\hat{\theta}^-}^-$ are the correlations which correspond to the breaking points.

To distinguish between direct comovement of risk measures, weak comovement and adverse comovement, (7) represents the pairwise and symmetric adjacencies. The first group contains pairs of SIFIs with a very strong positive dependence. The corresponding cells of A are coded with ones indicating active links in Fig. 2, i.e. $a_{ij} = 1$ (white) standing for highly positive

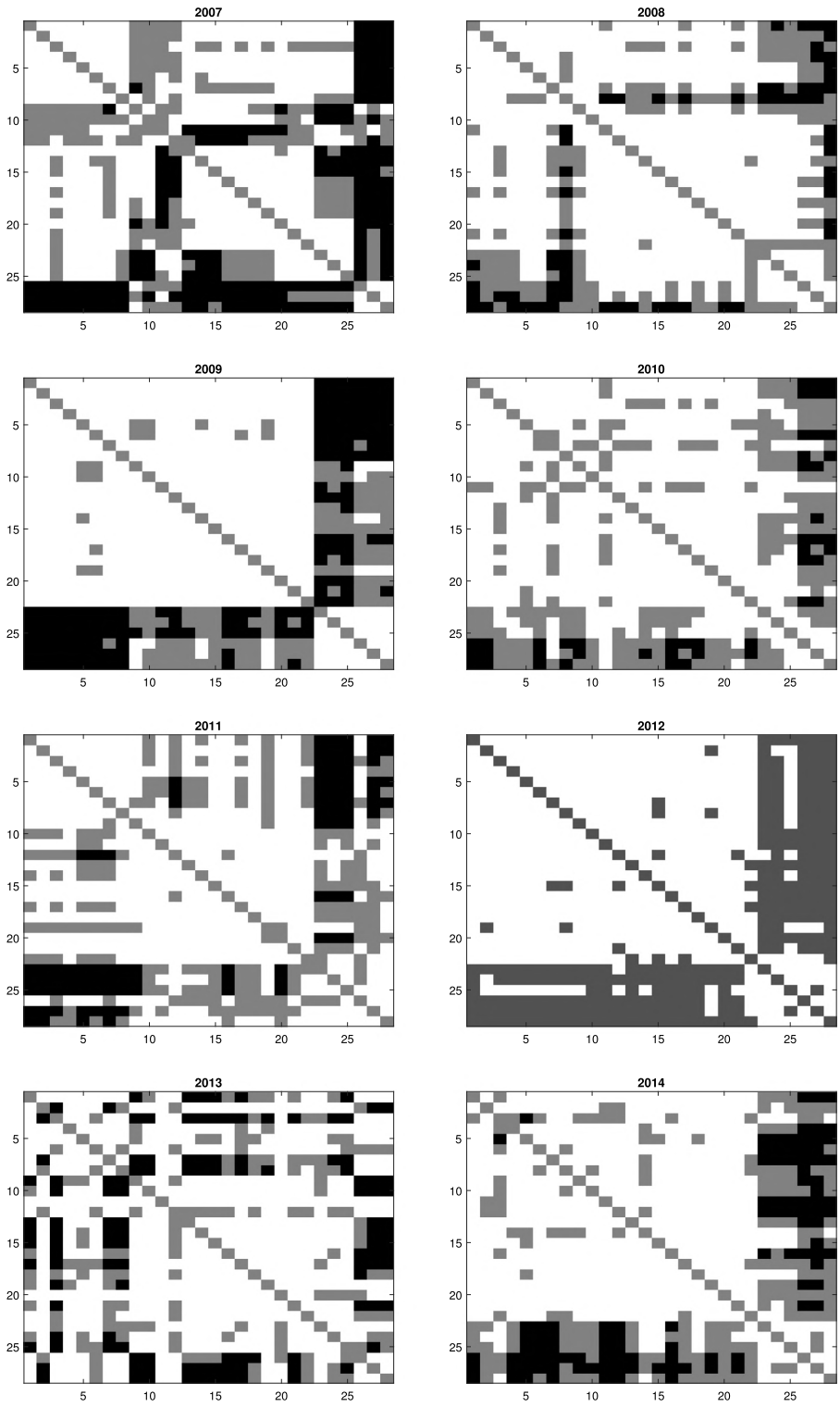

Fig. 2. Adjacency Matrix. Note: the highly positive correlation (in white), the highly negative correlation (in black) and the weak correlation (in gray) are shown in each of calender years. The diagonal cells show a trivial relation and in gray as a default color.

correlations. The 2nd group contains the pairs with highly negative correlation and is coded with $a_{ij} = -1$ (black). The elements of the adjacency matrix for the 3rd group are set to 0 (gray) reflecting weak or inactive linkages.

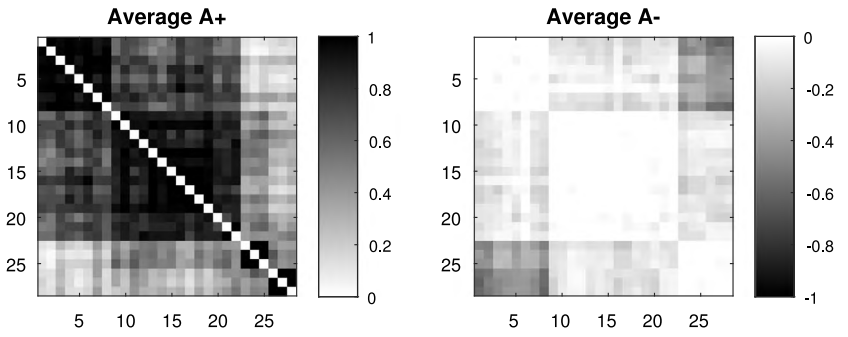

Fig. 3. The adjacency matrices A^+ and A^- averaged over time (2007–2015).

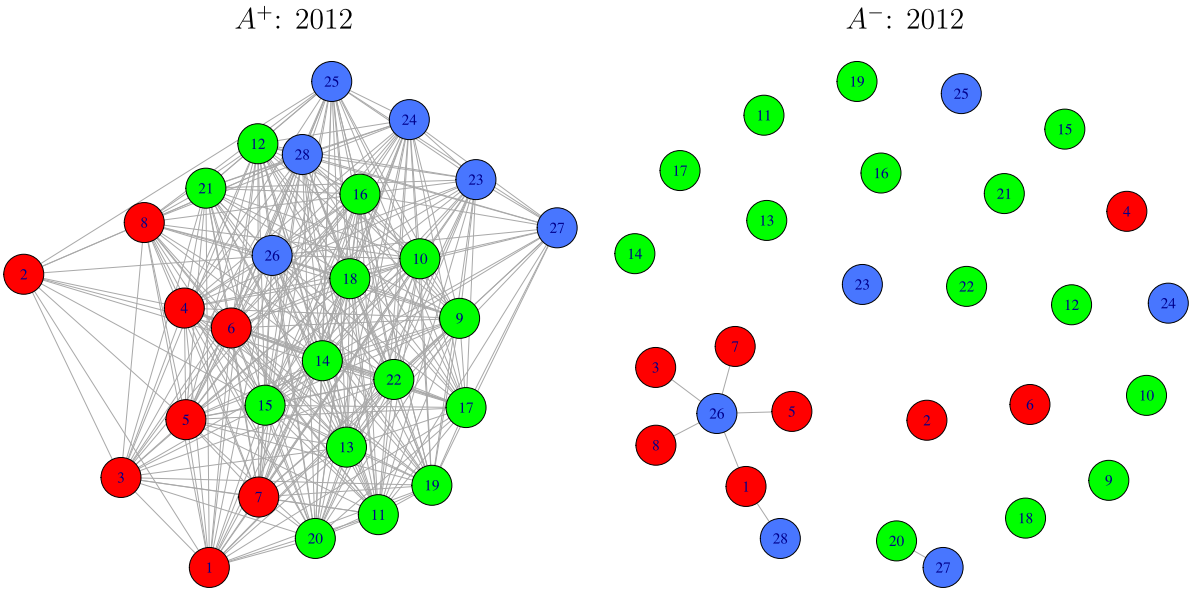

Fig. 4. Network visualization for A^+ (left) and A^- (right) for 2012. Green nodes indicate EU, red — US, and blue — the Asian SIFIs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

A sequence of snapshots can be merged into the averages of adjacency as shown in Fig. 3. Particularly, we separate the positive adjacencies into matrix A^+ from the negative into matrix A^-. One can therefore observe from the average of A^+ that the risk contagion emerges geographically in the form of four black blocks along the diagonal line, while the average of A^- illustrates that the Asian SIFIs are the major contributors of risk diversification in the world.

Note that the magnitudes of $\hat{\theta}^+$ and $\hat{\theta}^-$ are not necessarily symmetric. As the SIFIs are belonging the same sector, by nature they tend to move concurrently toward the same direction. In our investigation, a tendency of rising $\hat{\theta}^+$ in the crisis period along with a relatively lower n_2 (or higher n_1 since $n = n_1 + n_2$) is observed. For the case of $n_2 = 0$, the asymmetric breakpoint approach will coincide with the Ng (2006) approach.

The network in Fig. 4 visualizes one of the adjacency matrices of Fig. 2, e.g. the snapshot in 2012. We used an extended version of the force-directed layout suggested in Fruchterman and Reingold (1991) by minimizing the Euclidian distance between the nodes in the two networks. A highly connected A^+ network corresponds to a global contagion, whereas the sparse A^- network reveals scarce risk diversification. Interestingly, five of eight US SIFIs (node 1, 3, 5, 7, 8) are highly and negatively connected to Mitsubishi (node 26), showing a diversification benefit from Mitsubishi.

2.3.2. Testing the joint correlations in the positive and negative groups

The optimal break fractions determined in the previous section are of key importance for the construction of the adjacency matrix. The approach does not rely on the statistical properties of the spacing, but is mainly a clustering technique aimed to find the largest break in ordered positive or negative correlations. Further analysis of the spacings can reveal however additional insights into the correlation structure of the established groups from a statistical perspective. A natural question is to ask if the correlations $\rho^+_{v^+(i,j)} < \rho^+_{\hat{\theta}^+}$ and $\rho^-_{v^-(i,j)} > \rho^-_{\hat{\theta}^-}$ are simultaneously different from zero. If this is the case we can

conclude that the correlations exceeding the thresholds are also different from zero. A direct comparison of the correlations using spacings is not reasonable, since if all correlations are high and equal then the spacings are zero and render infeasible results. The following lemma summarizes several useful properties of the spacings (see Ng (2006), and Pyke (1965)).

Lemma 1. *Let $u_i \sim U[0,1]$ be iid for $i = 1, \ldots, n$. Let $\Delta_i = u_{[i]} - u_{[i-1]}$ be the ordered uniform spacings of order 1. Then $\mathsf{E}(\Delta_i) = 1/(n+1)$, $\mathsf{Var}(\Delta_i) = n/(n+1)^2(n+2)$ and $\mathsf{Cov}(\Delta_i, \Delta_j) = -1/(n+1)^2(n+2)$ implying $Corr(\Delta_i, \Delta_j) = -1/n$ for $i \neq j$.*

The joint distribution of spacings and further important properties can be found in Pyke (1965) and Arnold et al. (1992).

Let m_1 be the number of positive correlation between zero and $\rho^+_{\hat\theta^+}$ and m_2 the number of negative correlations between $\rho^-_{\hat\theta^-}$ and zero, i.e. $m_1 = \text{argmax}_k \{\rho^+_k < \rho^+_{\hat\theta^+}\}$ and similarly for m_2. Under the null of zero correlation we argue that the transformed correlations follow asymptotically half-uniform distribution with $\phi^+_i \sim U[0.5, 1]$ for $i = 1, \ldots, m_1$ and $\phi^-_i \sim U[0, 0.5]$ for $i = 1, \ldots, m_2$. Relying on Lemma 1 and noting that Δ^+_k are spacings on [0.5, 1] we obtain $\mathsf{E}(\Delta^+_k) = 0.5/(m_1+1)$ and $\mathsf{Var}(\Delta^+_k) = \frac{1}{4} m_1/(m_1+1)^2(m_1+2)$ and similarly for Δ^-_k. Moreover, the spacings are asymptotically uncorrelated. From this we can conclude that under the null of zero correlations it holds

$$\mathsf{Var}(\Delta^+_k + \Delta^+_{k-1})/2 = \mathsf{Var}(\Delta^+_k).$$

Therefore the variance ratio test of Cochrane (1988) is a convenient tool to test the autocorrelations of the spacings. The simplest version considers the ratio of the sample variance of 2-spacings $\Delta^+_k + \Delta^+_{k-1} = \phi^+_k - \phi^+_{k-2}$ to the sample variance of simple spacings.

$$\begin{aligned} SVR^+ &= \frac{\frac{1}{2}\widehat{\mathsf{Var}}(\Delta^+_k + \Delta^+_{k-1})}{\widehat{\mathsf{Var}}(\Delta^+_k)} - 1 \\ &= \frac{\frac{1}{2(m_1-3)} \sum_{j=2}^{m_1-1} \left\{ \Delta^+_k + \Delta^+_{k-1} - \frac{1}{m_1-2} \sum_{\ell=2}^{m_1-1} (\Delta^+_{(\ell)} + \Delta^+_{(\ell-1)}) \right\}^2}{\frac{1}{m_1-1} \sum_{k=1}^{m_1-1} \left\{ \Delta^+_k - \frac{1}{m_1-1} \sum_{\ell=2}^{m_1} \Delta^+_{(\ell)} \right\}^2} - 1. \end{aligned}$$

Similarly, we can define SVR^- with m_1 replaced by m_2 and Δ^+ by Δ^-.

Lemma 2. *Under the null $\rho^+_k = 0$ for $k = 1, \ldots, m_1$ and $\rho^-_k = 0$ for $k = 1, \ldots, m_2$ it holds that $\sqrt{m_1} SVR^+ \xrightarrow{\mathcal{L}} N(0,1)$ as $m_1 \to \infty$, $\sqrt{m_2} SVR^- \xrightarrow{\mathcal{L}} N(0,1)$ as $m_2 \to \infty$ and they are asymptotically uncorrelated.*

The uncorrelatedness follows from the asymptotically diminishing correlations in Lemma 1. Relying on the last lemma we can consider the Wald type test for the null hypothesis that m_1 smallest positive and m_2 largest negative correlations are zeros:

$$m_1 [SVR^+]^2 + m_2 [SVR^-]^2 \sim \chi^2_2. \tag{8}$$

If the hypothesis is rejected, we can conclude that the smaller as well as larger correlations are both different from zero. If the null hypothesis is not rejected, we can apply a similar test to positive correlations lying in $[\rho^+_{\hat\theta^+}, 1]$ and for negative correlations lying in $[-1, \rho^-_{\hat\theta^-}]$. Again rejecting the null here leads us to the conclusion that the larger correlations (either in $[\rho^+_{\hat\theta^+}, 1]$ or in $[-1, \rho^-_{\hat\theta^-}]$) are significant.

Fig. 5 displays the p-values of the Wald test in (8) for the S^+, S^- and the L^+, L^- sets, respectively. The p-values of S^+, S^- greater than 0.05 are the majority, showing that the small (in absolute value) correlations neither from the positive nor from the negative correlations reject the null hypothesis that they are jointly zero. A few exceptions are observed in the European crisis period, say 2012–2013.

A necessary investigation applied to the L^+, L^- sets reveals the statistical significance of risk profile correlations. The p-value constellations as given in Fig. 5 thus support our definition (7) of the adjacency structure.

2.4. Systemic risk score and risk decomposition

In social networks one might consider nodes/persons with many connections to be important or central. Several concepts exist that identify centrality e.g. eigenvector centrality, etc. In the situation that we study here though, one wants to identify nodes which contribute most to systemic risk. This is then informative for the supervisory authority for deciding capital requirements and systemic risk charges. The risk decomposition proposed below conveys information of the contribution of each SIFI, which can be ideally be seen as a complement of the bucket approach proposed by BCBS.

In addition to the notion of interconnectedness, node characteristics, e.g. size of bank or risk level of bank are also decisive. A bank with bigger market capitalization is more capable of offering interbank loans to other financial institutions, describing the situation of "too-big-to-fail". It is understood that the risk in a connected network is provoked by either the compromise level of nodes or the degree of their connectedness, or even both.

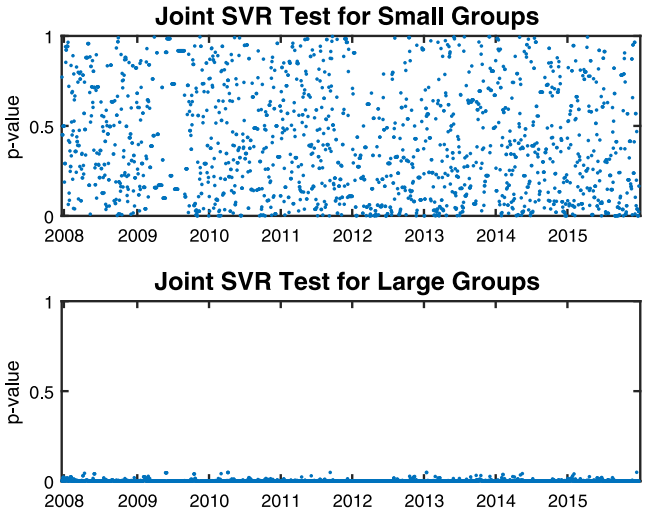

Fig. 5. p-values of the Wald test (8) for the correlations close to zero.

Motivated by Das (2016) who proposes a parsimonious way of quantifying aggregate risk in a network comprised of related entities, we combine adjacency matrix A together with a compromise loading (Fig. 2). The compromise level, e.g. the size, the capital allocation or characteristics used to measure the nodal risk (e.g. debt ratio, credit rating, probability of default, etc.) will be touching the issue of too-big-to-fail and Basel III regulation. Defining the level of compromises as nodal market capitalization $C = (C_1, \ldots, C_N)^\top \in \mathbb{R}^N$, we calculate the systemic risk score, S:

$$S(C, A) = C^\top AC \tag{9}$$

The aggregate risk score $S(C, A)$ is now decomposed into individual nodal contribution S_i, such that $S = \sum_{i=1}^N S_i$. $S(C, A)$ is linear homogeneous in C:

$$S = \sum_{i=1}^N S_i = \frac{\partial S}{\partial C_1} C_1 + \frac{\partial S}{\partial C_2} C_2 + \cdots + \frac{\partial S}{\partial C_N} C_N \tag{10}$$

The aggregate systemic risk score in (10) now allows ranking of nodal contributions. One can focus on the nodes with risk contributions arising from its connectedness since $\frac{\partial S}{\partial C_i} = 2\sum_{j=1}^N a_{ij} C_j$. Using (7), the contribution is positive (negative) if other nodes are positively (negatively) adjacent to node i. For a regulatory purpose, the node with a wide range of positive connectedness measured shall be prioritized and highly supervised as it has higher risk increment, while the nodes with more negative connectedness to others are very unlikely to trigger a system-wise risk.

Table 2 summarizes the results of risk decomposition of each node on annual basis ranging from 2007 to 2015. The aggregate systemic risk score rises in 2008 due to the outbreak of the US subprime crisis. During the EU crisis in 2012, one sees that the nodes 9–22 suffered in terms of their higher systemic risk contributions. The overall systemic risk score in a one year rolling window is shown in Fig. 6. Interestingly, $S(C, A)$ reaches several peaks as the number of negative correlations, n_2, turns to zero, indicating non-negative risk similarities among SIFIs and consequent higher systemic risk. The aggregate systemic risk score increased again in 2015 as the consequence of the Chinese stock market crash in summer 2015 caused by insolvent shadow banks, bursting asset bubbles and indebted local governments. Evidently, the aggregate risk score is an ideal measure to quantify overall systemic risk and its variations over time.

Through (10), one may attribute a high risk node as the one with a greater magnitude of compromise or high interconnectedness, or both. Such critical nodes need attention from regulators. One can e.g. find the node 11 (HSBC) as an important node in 2008, 2011 and 2013. To validate the results, we compare the identified important SIFIs in this study with the SIFIs assigned to Bucket 4 (highest loss absorbency group) reported in FSB Nov. 2015. We observe that the risk decomposition analysis yields a comparable list for this bucket but adds more quantitative insights.

A geographic analysis based on Table 2 documents to which extent systemic risk is attributed to a particular region where the headquarters of SIFIs are located. In general, the average scores in Europe are slightly larger than that of the US due to the fact that European banks build the majority of SIFIs. European banks are the largest geographic group within the 50 biggest global banks (21 banks from eight different European countries), while the US only contributes 7 banks. The Asian SIFIs "negatively" contribute to systemic risk score, which can be justified by their negative adjacencies with others, see Fig. 2. It documents that having Asian SIFIs in the system stabilizes the financial systems in terms of its role of risk diversification. However, this benefit becomes moderate during the distressed periods such as 2008, 2011.

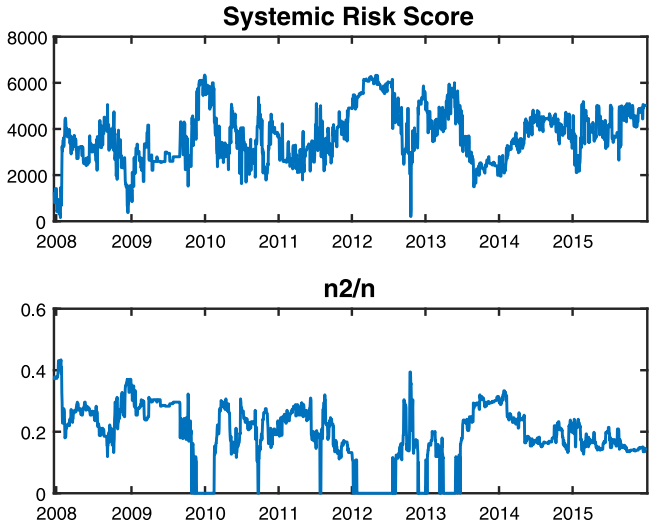

Fig. 6. Systemic risk score and ratio of negative correlation.

Table 2
Systemic risk decomposition. The largest systemic risk contributions are shown in red and the negative contributions in blue.

	SIFI	2007	2008	2009	2010	2011	2012	2013	2014	2015
1	JP MORGAN	153	189	134	152	98	189	79	153	98
2	BANK OF AMERICA	152	207	134	152	116	198	147	280	117
3	BANK OF NEW YORK MELLON	73	164	123	148	90	181	−10	115	74
4	CITIGROUP	152	197	133	206	152	197	215	143	162
5	GOLDMAN SACHS	147	208	94	147	69	191	173	87	51
6	MORGAN STANLEY	146	198	93	138	77	190	165	120	77
7	STATE STREET	96	161	128	121	56	169	22	104	80
8	WELLS FARGO	141	70	131	158	141	176	33	140	61
9	ROYAL BANK OF SCTL	38	209	191	145	163	199	84	154	108
10	BARCLAYS	55	227	182	209	165	190	120	163	108
11	HSBC	−11	228	172	145	201	191	246	118	101
12	STANDARD CHARTERED	16	212	152	179	121	178	126	101	103
13	BNP PARIBAS	108	219	182	202	192	182	121	172	110
14	CREDIT AGRICOLE	73	218	199	154	173	198	111	172	54
15	SOCIETE GENERALE	90	214	187	187	195	152	110	196	90
16	DEUTSCHE BANK	118	227	144	153	163	190	129	136	146
17	UNICREDIT	107	212	149	150	151	176	82	167	29
18	ING GROEP	115	223	168	169	195	186	153	205	144
19	SANTANDER	107	223	196	197	135	195	180	168	99
20	NORDEA BANK	119	207	154	207	146	181	198	181	63
21	CREDIT SUISSE GROUP N	148	209	157	219	201	183	133	174	104
22	UBS GROUP	159	185	158	159	169	203	187	169	114
23	BANK OF CHINA	51	161	−96	98	−43	63	180	37	108
24	ICBC	61	173	−70	135	−7	64	139	56	109
25	CHINA CON.BANK	42	189	−115	152	−25	151	102	19	117
26	MITSUBISHI UFJ	−162	182	20	−19	128	64	62	−72	0
27	MIZUHO	−133	135	−7	−28	29	54	24	−107	9
28	SUMITOMO.MITSUI	−149	−17	−24	−9	117	45	33	1	36
	Systemic Risk Score	2012	5232	3069	4028	3367	4536	3343	3253	2471
	Average score (US)	132	174	121	153	100	186	103	130	90
	Average score (Europe)	89	215	171	177	169	186	141	163	98
	Average score (Asia)	−148	100	−4	−19	91	54	40	−59	15

3. Tail event driven network quantile regression

The previous section provided a descriptive analysis of the SIFI and quantified the connectedness via adjacency matrices. In this section we present three issues on the network dynamics. First, it is of interest to relate the response of individual node at a given time point to the quantified connectedness at the previous time point. Second, it is very likely that the return of an individual SIFI asymmetrically responses to the network factors, that is, the node responses strongly to a network factor when it is under a stress (extremely negative returns) but may react mildly when it is experiencing an advance. The

likelihood of simultaneous slumps in the banking industry is potentially greater than that they boom together, which is a typical feature under systemic risk. Diebold and Yılmaz (2014) and Xu et al. (2016) find an increased total interconnectedness during the crisis period, resulting in more fragile financial markets evidence by high comovement, contagion and spillover. Recalling that systemic risk is defined through (i) large impact, particular in the downside constituting an asymmetric reaction; (ii) widespread coverage that can be examined through a panel or stack model; (iii) a ripple effect being detected by an intertemporal investigation, the TENQR model is proposed for accommodating three definitions. This aspect is beyond the work of Zhu et al. (2018) who focus on the asymptotics of the involved parameters.

Third, for financial assets, it is very likely that their returns are subject to the impact of neighboring assets in the sample industry or geographic region. Modeling returns for each asset individually revealed, as we show in the next section, interesting insights into the common behavior of assets from the same geographic region (see Brechmann et al. (2013). For this reason we consider the panel quantile regressions for the SIFIs from the US, Europe and Asia. By doing so, one may know which region is more vulnerable with respect to the network factor.

The extant research using the VAR framework requires ex-post analysis of the network structure and may also face the curse of dimensionality. This observation directs us towards a parsimonious model which incorporates tail sensitivity, allows for an asymmetric impact, establishes an intertemporal framework for investigating a ripple effect and evaluates regional reactions on network risk and the corresponding impulse response for the network shocks.

3.1. TENQR model

For the aforementioned purposes, we pay particular attention to tail events, like extremes or high level quantiles. A convenient framework which serves these objectives is the quantile regression. It yields insight into the interplay of the involved risk factors, and allows the investigator to explore a range of conditional quantile functions that link the dependent variable and the covariates in a continuous and smooth manner. However, this triggers problems if the number of covariates increases. Using a single index as in the work of Adrian and Brunnermeier (2016) is an alternative solution, it lacks though a systemic network insight as in Härdle et al. (2016). Besides, the financial network coherence is time-varying with different degree of interconnectedness over time. To tackle this we introduce network factors involving the adjacency matrices to represent the nodes in a financial system.

The enormous literature on asset pricing proposes numerous explanatory variables for the dynamics of SIFI returns. In the majority of the cases the model is restricted to linear regression on factors, like market-wide covariates, company-specific variables or autoregressive components. We partly follow this setting and involve the commonly used explanatory factors. The dynamics is reflected via the returns Y_{it} of SIFIs. The conditional quantile function of Y_{it} is

$$\begin{aligned} Q_{r,Y_{it}}(\tau) &= \gamma_{ri} + \beta_{r0}(\tau) + \beta_{r1}(\tau)Y_{i,t-1} + \beta_{r2}(\tau)Y_{i,t-2} + \sum_{\ell=1}^{L} \beta_{r\ell}^{(w)}(\tau)W_{\ell,t-1} \\ &\quad + \beta_r^+(\tau)\sum_{j=1}^{N} m_i^+(Y_{j,t-1}) + \beta_r^-(\tau)\sum_{j=1}^{N} m_i^-(Y_{j,t-1}) \quad \text{for } i \in \mathcal{R}_r, \end{aligned} \tag{11}$$

where $t = 1, \ldots, T$ and the sets $\mathcal{R}_r$ with $r = 1, 2, 3, 4$ contain the i-indices of SIFIs for each of the regions US ($r = 1$), Europe ($r = 2$), Asia ($r = 3$) or world ($r = 4$). $m_i^+(Y_{j,t-1})$ is a function of the positively linked neighboring nodes of node i, while $m_i^-(Y_{j,t-1})$ is a function of the negatively linked neighboring nodes of node i. The coefficients $\beta_r^+(\tau)$ and $\beta_r^-(\tau)$ respectively measure the impact of the positive and negative network factors and depend on τ only. $W_{\ell,t}$ represents the market-wide covariates. Here the volatility index VIX that serves as a proxy for the perceived market sentiment and the TED rate for the perceived credit risk are chosen.

The construction of m_i-function relies on the two adjacency matrices A^+ and A^- with the elements given as follows:

$$a_{ij}^+ = \begin{cases} 1, & \hat{\rho}_{v^+(i,j)} > \hat{\rho}_{\hat{\theta}^+}^+ \\ 0, & \text{else} \end{cases} \quad \text{and} \quad a_{ij}^- = \begin{cases} 1, & \hat{\rho}_{v^-(i,j)} < \hat{\rho}_{\hat{\theta}^-}^- \\ 0, & \text{else.} \end{cases} \tag{12}$$

Each row i of the matrix A^+ collects the SIFIs which exhibit strong positive correlation with the ith SIFI, while the corresponding row of the matrix A^- selects the SIFIs with strong negative dependence. The threshold values, $\hat{\rho}_{\hat{\theta}^+}^+$ and $\hat{\rho}_{\hat{\theta}^-}^-$ are defined and estimated in (5) and (6). This approach allows not only identifying for each time t the significant adjacencies but also disentangling the positive from the negative adjacencies. The resulting positive and negative *network factors* are defined as:

$$f_{i,t-1}^+ = \sum_{j=1}^{N} m_i^+(Y_{j,t-1}) = \frac{\sum_{j=1}^{N} a_{ij,t-1}^+ Y_{j,t-1}}{\sum_{j=1}^{N} a_{ij,t-1}^+}. \tag{13}$$

$$f_{i,t-1}^- = \sum_{j=1}^{N} m_i^-(Y_{j,t-1}) = \frac{\sum_{j=1}^{N} a_{ij,t-1}^- Y_{j,t-1}}{\sum_{j=1}^{N} a_{ij,t-1}^-}. \tag{14}$$

The factors f^+ and f^- in (13) and (14) measure the average impact from ith node. They quantify the average of its connected nodes in the system and specify their contribution to the next day return. Empirically, $f^+_{i,t-1}$ and $f^+_{i,t-1}$ are estimated at each time point by using a 250-day backward horizon.

In a financial perspective, the positive network factor measures the likelihood of triggering the risk contagion, whereas the negative network factor reflects the aim of risk diversification. The cutoff approach in (12) has advantages compared to full vector (quantile)-regressions since in larger networks a serious dimensionality problem surfaces. An extension to time-varying parameters is technically demanding and certainly requires specific assumptions on the data generating process. In the setup advocated here the relations within the network are treated as exogenous. The TENQR is estimated by minimizing:

$$\begin{aligned}
&\min_{\gamma_r,\beta_r(\tau)} \sum_{k=1}^{q}\sum_{t=1}^{T}\sum_{i\in\mathcal{R}_r} \\
&\rho_{\tau_k}\Big\{Y_{it} - \gamma_{ri} - \beta_{r0}(\tau_k) - \beta_{r1}(\tau_k)Y_{i,t-1} - \beta_{r2}(\tau_k)Y_{i,t-2} \\
&- \sum_{\ell=1}^{L}\beta^{(w)}_{r\ell}(\tau)W_{\ell,t-1} - \beta^+_r(\tau_k)f^+_{i,t-1} - \beta^-_r(\tau_k)f^-_{i,t-1}\Big\} \\
&+ \lambda\sum_{i\in\mathcal{R}_r}|\gamma_{ri}|,
\end{aligned} \tag{15}$$

where τ_k for $k = 1,\dots,q$ is a grid of the quantile levels of interest. It is possible to consider a weighted version of (15) if some quantiles are seen as more important for the analysis. Here $\rho_\tau(u) = u\cdot\{\tau - I(u<0)\}$ is an asymmetric loss function.

We use here a penalization in order to shrink the fixed-effects towards zero which was originally suggested by Koenker (2004). The approach is particularly advantageous in larger networks. Large values of λ render smaller fixed effects γ_{ri} and thus reduce the importance of node-specific information. On the other hand, small values might deteriorate the estimation results and inflate the impact of node-specific information. Lamarche (2010) suggests to select λ by minimizing the estimated asymptotic variance of the individual effects. The advantage of the approach is that it does not introduce a bias, but decreases the variance. Generally, λ varies with the number of nodes N, thus we can fix it in a finite network. Technically, the optimization is in line with the lasso-type linear regression, however, in a longitudinal quantile regression framework.

Algorithms to solve (15) can be found on www.quantlet.de and in Tran et al. (2016). The conditional quantile of Y_{it} is finally estimated by

$$\begin{aligned}
\hat{Q}_{r,Y_{it}}(\tau_k) &= \hat\gamma_{ri} + \hat\beta_{r0}(\tau_k) + \hat\beta_{r1}(\tau_k)Y_{i,t-1} + \hat\beta_{r2}(\tau_k)Y_{i,t-2} \\
&\quad + \sum_{\ell=1}^{L}\hat\beta^{(w)}_{r\ell}(\tau_k)W_{\ell,t-1} + \hat\beta^+_r(\tau_k)f^+_{i,t-1} + \hat\beta^-_r(\tau_k)f^-_{i,t-1}
\end{aligned} \tag{16}$$

for an appropriate value of $r = 1, 2, 3, 4$ and $k = 1,\dots,q$.

We do not consider any node-specific characteristics in model (11) for two reasons. First, typically such characteristics include the log firm size and the debt ratio computed as the total debt to assets ratio. However, these variables are computed at much lower frequency (mainly at annual rate) compared to the daily data on the asset returns. Second, the fixed effects γ_{ri} in the panel data model setup collect the sources of variability which are specific for each SIFI and, therefore, reflect the node-specific information. Note that γ_{ri}'s do not depend on the quantile level τ.

The role of the two autoregressive components is twofold. First, their objective is to capture the time-persistence of the returns. The lagged observations have typically little explanatory power, but are still popular factors in predictive models. Second, standard estimation of panel models with autoregressive part renders biased estimators both for linear models and in a quantile regression framework. Galvao (2011) tackles this problem and suggests to use $Y_{i,t-2}$ as an instrument aimed to reduce the bias in the effect of $Y_{i,t-1}$. Thus the 2nd order autoregressive element might be seen as a bias-decreasing instrument here. Note, however, that bias diminishes as T tends to infinity and is obviously a minor issue in our application. For this reason, we do not employ the two-stage technique of Galvao (2011).

3.2. Estimation results

To estimate the above model we collect daily data on the SIFI returns for the period from 01.01.2007 till 31.12.2015. The returns Y_{it} exhibit heavy tails; Fig. 7 shows the normal QQ plots of the returns for Citigroup (in the US), Mitsubishi (in Asia) and HSBC (in Europe), the thick tail distributional property is documented. Let us focus more on the tails: to compare the distribution of the SIFI returns to the network factors we plot the corresponding QQ-plots in Fig. 8. For Citigroup both factors possess lighter tails than the SIFI returns, while for the other two companies we have opposite evidence. Particularly, the negative factor has much heavier tails for EU and Asian SIFIs. To assess the relationship between the extremes and the network factors we consider the cross-quantilogram suggested by Han et al. (2016).

$$\varrho_i(\tau_1,\tau_2) = \frac{\mathsf{E}[\psi_{\tau_1}(Y_{it} - F_i^{-1}(\tau_1))\psi_{\tau_2}(f_{t-1} - F^{-1}(\tau_2))]}{\sqrt{\mathsf{E}[\psi^2_{\tau_1}(Y_{it} - F_i^{-1}(\tau_1))]\cdot\mathsf{E}[\psi^2_{\tau_2}(f_{t-1} - F^{-1}(\tau_2))]}},$$

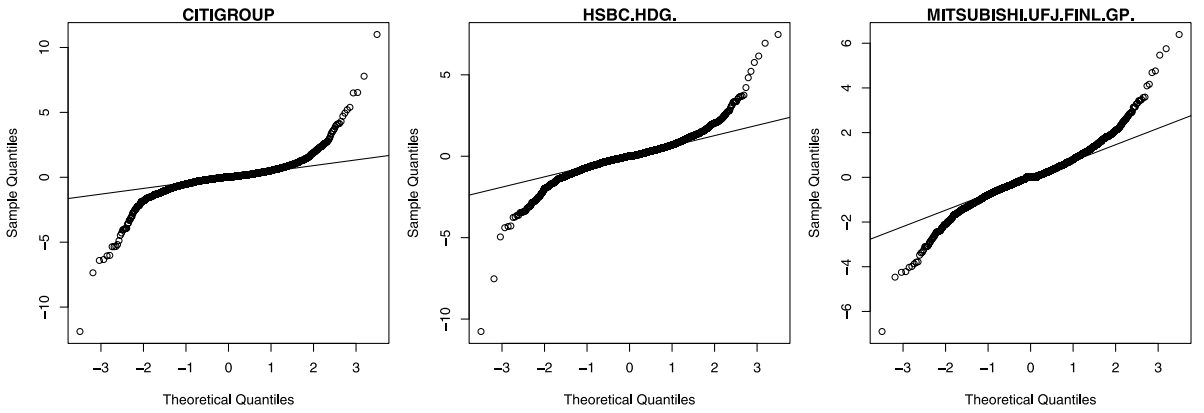

Fig. 7. QQ plots for the returns of Citigroup, HSBC and Mitsubishi Financial and the Gaussian distribution (full sample estimation).

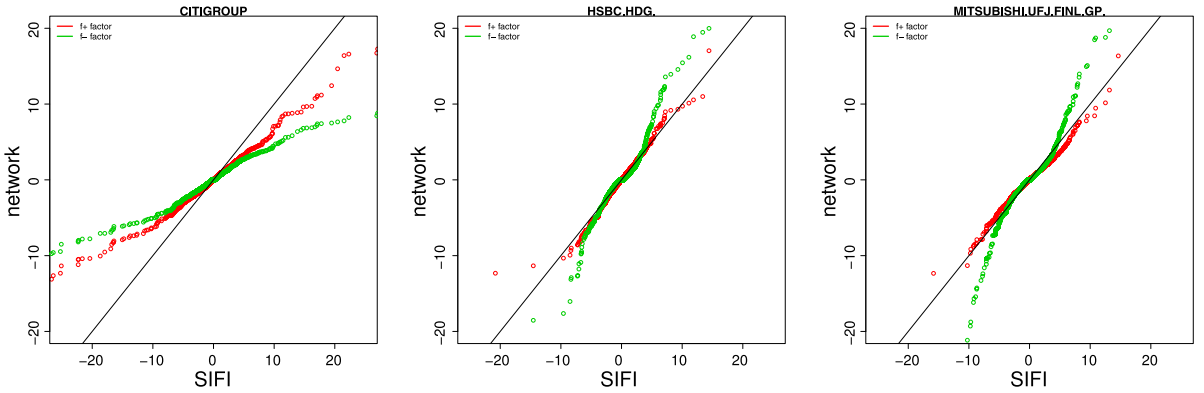

Fig. 8. QQ plots of f_t^+ and f_t^- vs. the returns of Citigroup, HSBC and Mitsubishi Financial.

where f is one of the network factors f^+ or f^- and F^{-1} is its quantile function. F_i^{-1} denotes the quantile function of the ith SIFI return and $\psi_a(u) = I\{u < 0\} - a$. The cross-quantilogram can be estimated by replacing the quantile functions with their sample counterparts. We calculate the first order cross-quantilogram between Y_{it} and the network factors based on the regional set of the SIFIs. Fig. 9 shows the correlation of different quantiles of the returns for the three selected SIFIs with the 10%, 50% and 90% quantiles of the network factors defined in (13) and (14). Consider first the lower values of τ_1 which correspond to highly negative returns. For Citigroup and HSBC we observe that these returns react more strongly to the lower quantiles of the negative network factor than of the positive network factor. This is intuitively correct, since the negative network factor f^- cumulates turbulent market situations and renders even smaller SIFI returns. Note that for the negative factor the curves for $\tau_2 = 0.5$ and $\tau_2 = 0.9$ are very close implying that the upper quantiles of this factor do not influence the SIFIs under stress, which is again an economically reasonable observation. If we move to higher τ_1 values we observe an inverse situation with the positive factor becoming more relevant at high quantiles. For the Asian companies, for example for Mitsubishi, the evidence is mixed. This will be supported below by the results from the quantile regression.

3.2.1. The results from the TENQR

The geographic regions of the SIFIs provide additional information. Given the network factors in (13) and (14) we select only the dependent variable for SIFIs from a specific region. As before we compare the sensitivity and predictability of these two network factors at multiple percentiles. By comparing the network effect curves from the geographic regions one may shed some light on the vulnerability in a geographic perspective.

Let us first concentrate on the positive network factor. Asia displays a U-shape network coefficient curve, indicating a comparable impact of the network effect on the right- and left-tail of return distribution but a relatively moderate effect on the center of distribution. However, the EU and the US SIFIs exhibit a monotonic decreasing coefficient curve. In addition, the Asian SIFIs exhibit the highest sensitivity to the network factor among the three regions. The reason is that for $r = 3$, the Asian region, the active cells in $f_{i,t}^+$ are exclusively concentrating on their Asian neighbors. Referring to Figs. 2 and 3, the Asian SIFIs move tightly together with their Asian neighbors and are different. By contrast, the geographic effect in the EU or the US is not as strong as in Asia because the positive network effect not only contributes geographically but globally. As can be seen in Fig. 2, the US SIFIs and the EU SIFIs entail the across-border risk adjacencies.

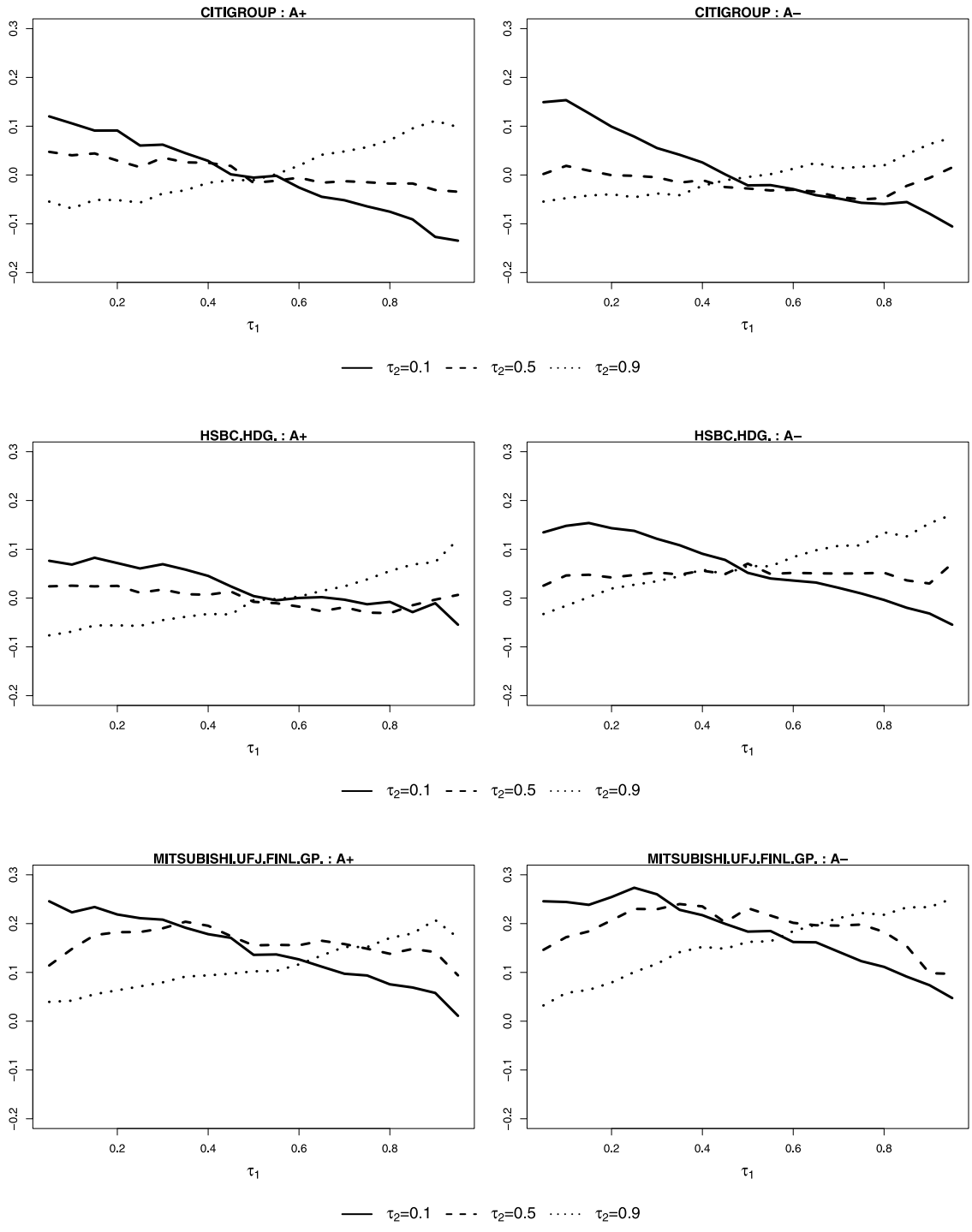

Fig. 9. Quantilograms of the return of Citigroup, HSBC and Mitsubishi Financial with the 10%, 50% and 90% quantiles of network factor (full sample estimation).

For the US and EU, the monotonically downward curves indicate an asymmetric response to the positive network factor (with a small exception at $\tau = 0.1$). The network factor contributes to the return prediction in the lower quantile of return distribution more than it can create under the higher quantile situation, showing an asymmetric impact of the positive network factor over τ values. Economically speaking, the network risk contagion tends to happen at the lower quantile

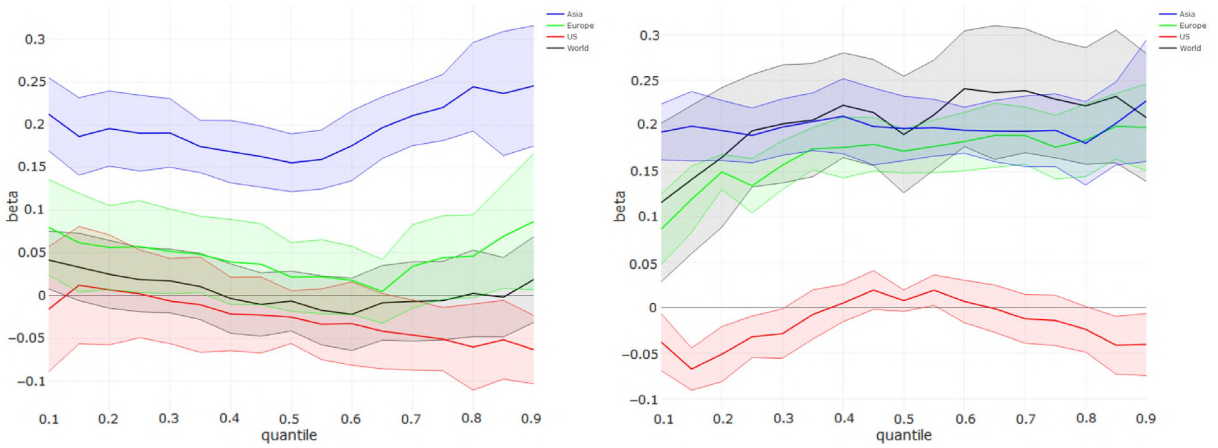

Fig. 10. Slopes from the quantile regressions of SIFI returns from different geographic regions on the network factor with adjacency matrices A^+ (left) and A^- (right). The colored area shows the 95% pointwise confidence intervals (full sample estimation). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

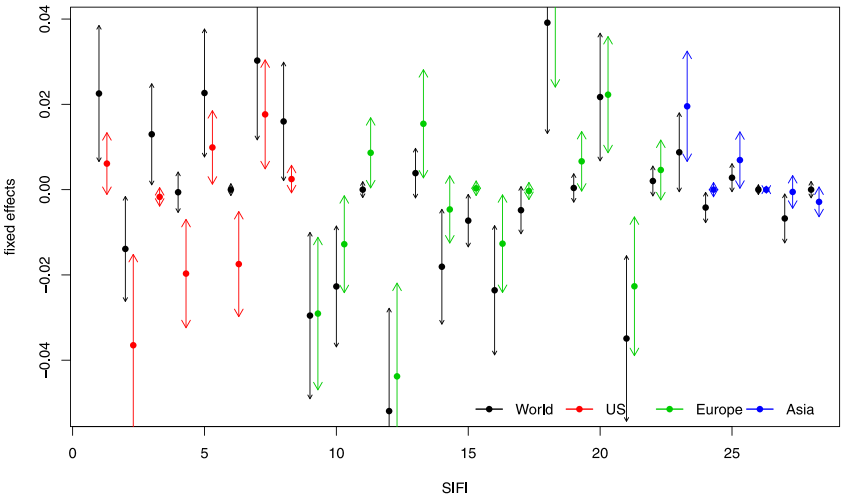

Fig. 11. Fixed effects from different geographic regions (full sample estimation) with 95% confidence intervals.

levels. Such an asymmetric impact has also been documented in research on comovement, contagion and spillover in crisis periods, see Dungey and Gajurel (2015) and Diebold and Yılmaz (2014).

Let us now turn to f^-. Recall that the A^- matrix collects the adverse risk movements. Having such adverse risk movements in the banking system alleviates the ripple effect and stabilizes its functioning. Surprisingly, the U-shape characteristic in Asian is no longer existing, but still has a relatively high response curve as compared to the US. This reflects that the Asian SIFIs are dominating elements in the A^- matrix. They are benefiting of higher risk diversification and this benefit is invariant over τ.

The quantile curves in the EU and world shift up and exhibit monotonic increasing behavior. From a financial perspective, this may be interpreted as risk diversification and it moves along the τ-axis from right to left, showing that the diversification gain monotonically decreases. It is understood that during market downturn, the risk profile similarities ascend, resulting in a limited diversification benefit. The diversification gain in the US SIFIs is very limited since the Asia SIFIs are their exclusive diversification contributors. The 95% confidence bands in Fig. 10 are actually connected confidence intervals. We took the liberty to address these as bands though since the asymptotic difference is minor. First, the impact of the network factor significantly differs from one geographic region to another. Second, the zero value is not covered by the confidence bands indicating a significant impact factor.

The fixed effect in (11) are documented in Fig. 11. The individual specific effect is very prominent in the US SIFIs with higher variations in $\hat{\gamma}_{ri}$. The heterogeneity among the EU SIFIs is also well evident. By contrast, the Asian SIFIs have less heterogeneity, implying that a relatively milder individual effect is attributed to the less flexible financial polices in Asia or

less diversity among banks. The state-owned banks are the majority in China, and their business types somehow are quite uniform.

In a nutshell, the TENQR method documents the importance of the network factors on the tail distribution, and specify the distinct impacts from the positive and the negative network factor. The implied vulnerabilities across regions are in accordance with the results of systemic risk decomposition. Having these efforts, we contribute to a "manageable" systemic risk. The supervisors are able to identify the central SIFIs with higher risk contributions, to measure the resulting connectedness in a system, and to evaluate the impact of network on the conditional quantile of a response.

4. Conclusion

There is no doubt that systemic risk depends on the interdependence and the joint dynamics of SIFIs in stress situations. Such tail event driven stress on the financial system is conveniently modeled in a network topology. The possible overstretching of positions i.e. the likeliness of common tail events is addressed in a panel quantile regression framework with fixed effect. The choice of the adjacency matrix is based on the similarity of risk profiles. We propose an asymmetric breakpoint approach to group pairwise correlations into strongly positive, strongly negative and weak ones. A joint SVR test confirms the competence of the breakpoint approach and the design of adjacency matrices. Employing the positive and negative network factors, we implement TENQR (Tail Event Network Quantile Regression) in order to capture risk propagation/diversification in the regional level but also in the global level. The network with adjacency information allows quantifying the risk contribution of each SIFI.

The network quantile curves vary between the positive v.s. the negative network factor, and change from one region to another one. The risk contagion effect is concentrating on lower quantiles, whereas the diversification gain is more tangible at higher quantiles. The TENQR model highlights the network factors and thereby allows investigating the joint dynamics in a stress situation of the financial system. The methodologies and quantitative techniques in this study are tailored for network analysis of systemic risk.

Acknowledgments

Financial support from the German Research Foundation (DFG) via Collaborative Research Center 649 "Economic Risk" and International Research Training Group 1792 "High Dimensional Nonstationary Time Series", Humboldt-Universität zu Berlin, is gratefully acknowledged. We would like to thank Francis Diebold, Simone Manganelli and Kamil Yilmaz for helpful suggestions.

References

Adrian, T., Brunnermeier, M.K., 2016. CoVaR. Amer. Econ. Rev. 106, 1705–1741.
Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., 1992. A First Course in Order Statistics. John Wiley & Sons.
Barigozzi, M., Brownlees, C.T., 2015. NETS: Network Estimation for Time Series, Manuscript.
Bluhm, M., Krahnen, J.P., 2014. Systemic risk in an interconnected banking system with endogenous asset markets. J. Financ. Stability 13, 75–94.
Borak, S., Härdle, W., López-Cabrera, B., 2013. Statistics of Financial Markets. Springer.
Brechmann, E., Hendrich, K., Czado, C., 2013. Conditional copula simulation for systemic risk stress testing. Insurance Math. Econom. 53, 722–732.
Cochrane, J.H., 1988. How big is the random walk in GNP?. J. Political Econom. 96 (5), 893–920.
Das, S.R., 2016. Matrix metrics: Network-based systemic risk scoring. J. Altern. Invest.: Special Issue on Systemic Risk 18 (4), 41–55.
Diebold, F.X., Yılmaz, K., 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. J. Econometrics 182 (1), 119–134.
Dungey, M., Gajurel, D., 2015. Contagion and banking crisis: International evidence for 2007-2009. J. Bank. Finance 60, 271–283.
Fan, Y., Härdle, W.K., Wang, W., Zhu, L., 2016. Single index based CoVaR with very high dimensional covariates. J. Bus. Econ. Stat. http://dx.doi.org/10.1080/07350015.2016.1180990.
Fruchterman, T., Reingold, E., 1991. Graph drawing by force-directed placement. Softw. Pract. Exp. 21 (1), 1129–1164.
Galvao, A.F., 2011. Quantile regression for dynamic panel data with fixed effects. J. Econometrics 164, 142–157.
Han, H., Linton, O., Okac, T., Whang, Y.-J., 2016. The cross-quantilogram: Measuring quantile dependence and testing directional predictability between time series. J. Econometrics 193 (1), 251–270.
Härdle, W.K., Wang, W., Yu, L., 2016. TENET: Tail-Event driven NETwork risk. J. Econometrics 192 (2), 499–513.
Hautsch, N., Schaumburg, J., Schienle, M., 2014. Financial network systemic risk contributions. Rev. Finance 19 (2), 685–738.
Koenker, R., 2004. Quantile regression for longitudial data. J. Multivariate Anal. 91, 74–89.
Laeven, L., Ratnovski, L., Tong, H., 2015. Bank size, capital, and systemic risk: Some international evidence. J. Bank. Finance 192 (2), 499–513.
Lamarche, C., 2010. Robust penalized quantile regression estimation for panel data. J. Econometrics 157, 396–408.
Ng, S., 2006. Testing cross-section correlation in panel data using spacings. J. Bus. Econom. Statist. 24 (1), 12–23.
Pyke, R., 1965. Spacings. J. R. Stat. Soc. Ser. B Stat. Methodol. 27, 395–449.
Tran, N.M., Burdejová, P., Osipenko, M., Härdle, W.K., 2016. Principal Component Analysis in an Asymmetric Norm. SFB 649 Discussion Paper 2016-040, HU Berlin.
Xu, X., Chen, C.Y.-H., Härdle, W.K., 2016. Dynamic Credit Default Swaps Curve in a Network Topology. SFB 649 Discussion Paper 2016-059, HU Berlin.
Zhu, X., Pan, R., Li, G., Liu, Y., Wang, H., 2017. Network vector autoregression. Ann. Statist. 45, 1096–1123.
Zhu, X., Wang, W., Wang, H., Härdle, W., 2018. Network quantile autoregression. J. Econ. (forthcoming).