Luu, Duc Thi

Portfolio Correlations in the Bank-Firm Credit Market of Japan

Computational Economics

Provided in Cooperation with:
Springer Nature

Suggested Citation: Luu, Duc Thi (2021) : Portfolio Correlations in the Bank-Firm Credit Market of Japan, Computational Economics, ISSN 1572-9974, Springer US, New York, NY, Vol. 60, Iss. 2, pp. 529-569,
https://doi.org/10.1007/s10614-021-10157-y

Computational Economics (2022) 60:529–569
https://doi.org/10.1007/s10614-021-10157-y

Portfolio Correlations in the Bank-Firm Credit Market of Japan

Duc Thi Luu[1]

Accepted: 15 July 2021 / Published online: 31 July 2021

Abstract

The recent global financial crisis has shown portfolio correlations between agents as one of the major channels of risk contagion and amplification. In this work, we analyse the structure and dynamics of the cross-correlation matrix of banks' loan portfolios in the yearly bank-firm credit network of Japan during the period from 1980 to 2012. Using the methods of Random Matrix Theory (RMT), Principal Component Analysis and complex networks, we aim to detect non-random patterns in the empirical cross-correlations as well as to identify different states of such correlations over time. Our findings suggest that although a majority of portfolio correlations between banks in lending relations to firms are contributed by noise, the top largest eigenvalues always deviate from the random bulk explained by RMT, indicating the presence of non-random patterns governing the correlation dynamics. In particular, we show that this dynamics is mainly driven by a global common factor and a couple of "groups" factors. Furthermore, different states in the credit market can be identified based on the evolution of eigenvalues and associated eigenvectors. For example, during the asset price bubble period in Japan from 1986 to 1991, we find that banks' loan portfolios tend to be more correlated, showing a significant increase in the level of systemic risk in the credit market. In addition, building Planar Maximally Filtered Graphs from the correlations of different eigenmodes, notably, we observe that the local interaction structure between banks changes in different periods. Typically, when the dominance of a group of banks in one period gradually vanishes, the credit market starts to build-up a different structure in the next period in which another group of banks will become the main actors in the backbone of the cross-correlations.

Keywords Bank lending · Portfolio correlations · Systemic risk · Random matrix theory · Principal component analysis · Correlation-based filtered methods

✉ Duc Thi Luu
d.t.luu@economics.uni-kiel.de

[1] Institute of Economics, University of Kiel, Kiel, Germany

1 Introduction

Over the last few years, catastrophic cascade of failures in interdependent systems has received a remarkable attention in network science. Findings in this line of research show that the robustness of a system crucially depends on both its internal structure as well as its pattern of relations to other interdependent ones (e.g. Smart et al. 2008; Buldyrev et al. 2010; Huang et al. 2011; Brummitt et al. 2012; Reis et al. 2014; Liu et al. 2016; Bianconi 2018).

In the banking system, the recent global financial turmoil has demonstrated that besides direct credit linkages[1], indirect relations between banks are also the relevant channels of risk contagion and amplification. For instance, overlapping portfolios due to common asset holdings have been widely considered as one of the major sources of risk contagion (e.g. May and Arinaminpathy 2010; Beale et al. 2011; Huang et al. 2013; Caccioli et al. 2014, 2015; Greenwood et al. 2015; Lillo and Pirino 2015; Levy-Carciente et al. 2015; Corsi et al. 2016). Analogously, credit relationships between banks and firms have been suggested as another important channel of the propagation of distress between the financial system and the real sector of the economy (e.g. Aoyama et al. 2013; Aoyama 2014; de Castro Miranda and Tabak 2013; Lux 2016). Accordingly, since banks can be indirectly linked through a set of joint exposures to firms (i.e. due to loan portfolio overlaps), the distress originating from a group of banks or firms can be propagated through the whole system. For these reasons, a number of empirical studies have been devoted to uncovering the complexity of the topological structure of bank-firm credit markets world-wide (e.g. Fujiwara et al. 2009; De Masi et al. 2011; De Masi and Gallegati 2012; Fricke 2016; Marotta et al. 2015, 2016; Fricke and Roukny 2020; Luu and Lux 2018, 2019; Lux 2020).

It is worthwhile to emphasize that besides portfolio overlaps, portfolio correlations also play a crucial role in risk management and have been seen as one of the major sources of financial instability. From a perspective of bankers, the understanding of the structure of correlations (or co-movements) between banks in providing loans to non-financial firms as well as the identification of factors governing such correlations are essential for managing credit risk (e.g. Bülbül and Lambert 2012; Fenech et al. 2015). The systematic factors driving the correlations represent market risks that cannot be diversified a way. In contrast, idiosyncratic risks associated with non-systematic components can be diversifiable as the dimension of the system grows (e.g. Chamberlain and Rothschild 1983). Furthermore, from a macroeconomic perspective, in a banking system where bank portfolios are highly unified or coupled, a common shock might trigger the entire system to follow the similar adjustment strategies. Even if the original shock is relatively small, it would result in a significant change in the credit supply to the other sectors of the economy, which in turn might lead to a large impact on macroeconomic fluctuations (e.g. Kiyotaki and Moore 1997; Khwaja and Mian 2008; Gertler and Kiyotaki 2010; Jorda et al. 2013; Bassett et al. 2014;

[1] See, for example, Allen and Gale (2000), Nier et al. (2007), Haldane and May (2011), Acemoglu et al. (2015), among others.

Brunnermeier and Schnabel 2016; Cingano et al. 2016; Gambetti and Musso 2017; Bentolila et al. 2018; Azariadis 2018; Alfaro et al. 2018; Amiti and Weinstein 2018). However, so far, the structure and the key drivers of correlations between banks in lending to firms still remain a puzzling mystery.

This study, therefore, aims to contribute to a deeper understanding of the structure and the dynamics of the cross-correlation matrix of banks' loan portfolios. In particular, our main objectives are to answer the following research questions that have not been intensively studied in related literature: (i) how to detect non-random patterns and latent factors (if they exist) in the empirical cross-correlation matrix of banks' loan portfolios?; and (ii) how to identify different states in the correlation dynamics over time?; and (iii) how to quantify systemic risk from loan portfolio correlations in the credit markets?.

To guide our analysis, we employ the methods of Random Matrix Theory (RMT) and Principal Component Analysis (PCA). The methods of RMT have been widely used to filter noise as well as to extract latent information embedded in empirical correlations for time series data on stock prices/returns (e.g. Laloux et al. 1999, 2000; Plerou et al. 2000, 2002; Kim and Jeong 2005; Bouchaud and Potters 2009; Wang et al. 2011; Meng et al. 2014; Jiang et al. 2014; Uechi et al. 2015; MacMahon and Garlaschelli 2015), on financial transactions (e.g. Kyriakopoulos et al. 2009; Fricke 2012), etc. Empirical evidence also shows that RMT is a useful tool to enhance the Markowitz mean-variance optimization procedure (e.g. Plerou et al. 2000; Laloux et al. 2000; Sharifi et al. 2004; Daly et al. 2008; Bai et al. 2009; Quintana et al. 2015). Recently, RMT has been used to study synchronizations between different macroeconomic indicators (e.g. Ormerod and Mounfield 2000; Ormerod 2008; Iyetomi et al. 2011; Yoshikawa et al. 2015; Lux et al. 2020). Indeed, the applications of RMT to the spectral analysis of various types of complex networks have also received considerable attention in related literature (Bandyopadhyay and Jalan 2007; de Carvalho et al. 2009; Potestio et al. 2009; Jalan 2009; Tran et al. 2013; Dumitriu and Johnson 2016), but to the best of our knowledge none of them has studied the spectral properties of cross-correlations between nodes in one set in the formation of links or the strength of interactions with nodes in another set (e.g. like the correlations between banks in providing loans to firms in the other sectors of the economy as in the present study).

Practically, in this study, we will, first, compare the eigenvalue spectrum of the empirical cross-correlation matrix with the theoretical one implied by RMT. Second, focusing on a group of the largest eigenvalues significantly deviating from random bulk predicted by RMT, we extract latent information and infer the significant factors governing the correlation dynamics. This approach allows us to have a proper selection of the strongest factors containing genuine information about the properties of the observed data that are different from what would be expected under a null model. In the next step, in order to identify different states in the credit market, we examine the evolution of the largest eigenvalues and their corresponding eigenvectors over time. In addition, building Planar Maximally Filtered Graphs (PMFG) (Tumminello et al. 2005; Pozzi et al. 2008; Di Matteo et al. 2010) from the correlation structure for different eigenmodes, we also analyse the network structure of correlations and investigate how this network changes in

different periods. Furthermore, to infer systemic risk arising from credit extensions from banks to firms, we measure the absorption ratios that capture the fraction of the total variance explained by a finite number of the first principal components (Kritzman et al. 2011; Billio et al. 2012; Zheng et al. 2012; Lux et al. 2020). It should be emphasized that the procedure used in the present paper can be employed to filter noise signals and extract significant patterns out of different large bipartite networks including those in economics and finance such as bank-borrower credit networks, investor-asset networks.

Perhaps, among different studies on the bank-firm credit relations, the work of Fujiwara et al. (2009) is closest to our study. Based on eigenvalue analysis, Fujiwara et al. (2009) measure the fragility scores in terms of dynamical propagation of influences from lenders to borrowers and vice versa. However, there exist important distinctions between ours and theirs. To name a few, first, it should be emphasized that the main difference to the present paper is, instead of analysing the correlations between banks in lending to firms, Fujiwara et al. (2009) define a new weighted matrix P capturing the accumulated interactions between the two sectors. Hence, their approach cannot answer whether the loan portfolios of banks tend to be more/less correlated in the credit market in different periods. Furthermore, in such a setting, the largest eigenvalue is always equal to 1, and few top largest eigenvalues (i.e. the second and the next largest eigenvalues of P) are important in the propagation of influences from banks to firms and vice versa over a finite time-scale.[2] Different from their approach, we focus on the largest eigenvalues that significantly deviate from the spectral distribution explained by an appropriate null model (e.g. the distribution given by the so-called Marchenko-Pastur law (Marčenko and Pastur 1967; Tao and Vu 2012)).[3] Such eigenvalues and the corresponding eigenvectors are analysed as they might contain the genuine information of particular economic mechanisms responsible for an unexpectedly high level of correlations between banks in lending to the non-financial sector of the economy.

The remainder of this paper is structured as follows. Section 2 briefly describes the data and methods used to analyse portfolio correlations. Section 3 reports the main findings for the weighted version of the bank-firm credit network of Japan. Discussions and concluding remarks are in Sect. 4. In the "Appendix", we provide additional results for portfolio overlaps and similarities, the fractions of eigenvalues in different ranges, the distributions as well as the evolution of the eigenvector

[2] In their study, briefly, from the matrix of weights W^{bf} representing loans from banks to firms defined in the next part of the present study, they obtain a matrix A that stands for relative amounts of lending by banks to firms and a matrix B that stands for the relative amount of borrowing by firms from banks. Then the matrix $P = AB$ will capture the accumulated reflections and influences between banks and firms. Assume that $\lambda_1^{(P)} \geq \lambda_2^{(P)} \geq ... \geq \lambda_N^{(P)}$ are the eigenvalues of P, it can be shown that $\lambda_1^{(P)} = 1$. Therefore, over a finite time-scale, only few next largest eigenvalues $\lambda_2^{(P)}$... are important in the dynamics of the propagation of influences from banks to firms and vice versa. For more details, we refer readers to the study of Fujiwara et al. (2009).

[3] Note that to deal with potential effects of the presence of extreme loan values on the largest eigenvalues, we also consider the other upper bounds obtained from heavy-tailed random matrices (see Sect. 2.2.2).

elements, and the more detailed correlation network structure between banks based on the method of PMFG. We also summarize the main results for the binary version of the credit network.

2 Data and Methods

2.1 Data

We use a yearly data set containing credit relations between banks and large firms in Japan over the period from 1980 to 2012. The data were obtained from the firms' financial statements and the survey by Nikkei Media Marketing, Inc. in Tokyo (commercially available).[4] This covers the most critical time and events over the past few decades in the Japanese economy, e.g. the bubble period from 1986 to 1991, the burst of the bubble economy and the stagnation time in the subsequent years (the so called "lost decade") (e.g. see Hoshi and Kashyap 2000; Brunnermeier and Schnabel 2016; Himino 2021), the recent global financial crisis.[5]

Based on the maturity of loans, we separately consider three layers of lending relationships from banks to firms, i.e. total lending relationships (namely layer 1), short-term lending relationships (namely layer 2), and long-term lending relationships (namely layer 3). Note that the layer 2 captures all loan contracts lasting less than or equal to one year. In contrast, the layer 3 only consists of loan contracts that exceed one year.

For the sake of conciseness, we exclude other financial institutions including insurance companies and a small number of aggregated banks.[6] In addition, it should be emphasized that for each maturity-based layer, we focus on the aggregate amount of loans from each bank to each firm at the end of the fiscal year, meaning that we ignore contract-specific frequencies (i.e. separate times that a bank lends to a firm during the fiscal year).

2.2 Methods

2.2.1 Representation of a Bank-Firm Network, Portfolio Overlaps and Projection Matrices

We shall start with some basic notations and definitions of a bank-firm credit network. Assume that in each year we have a weighted bipartite network consisting of N banks lending to N_F firms. Let $W^{bf} = \{w_{ij}^{bf}\}_{N \times N_F}$ be the matrix of weights (i runs from 1 to N standing for N banks, and j runs from 1 to N_F representing N_F

[4] For a more detailed explanation about the data set, we refer readers to, for example, Fujiwara et al. (2009), Marotta et al. (2015), Marotta et al. (2016).

[5] Although additional analysis for more recent data could provide more insights on the evolution of the portfolio correlations among banks in lending to non-financial firms, we leave it for future study once data is available.

[6] Each aggregated bank consists of a group of small/local banks. However, the main conclusions do not change very much when we include insurance companies and aggregated banks in the data, since the market shares of these companies and banks are relatively small.

firms). Each element w_{ij}^{bf} is non-negative and implies the monetary amount of loan[7] that bank i lends to firm j. Notice that from the matrix W^{bf}, we obtain the adjacency matrix $A^{bf} = \{a_{ij}^{bf}\}_{NxN_F}$, where $a_{ij}^{bf} = 1$ if $w_{ij}^{bf} > 0$, and zero otherwise. This adjacency matrix is associated with the binary version of the bank-firm credit network.

From a perspective of network theory, one can measure the structural similarity between banks' loan portfolios based on various measures of the similarity between two vectors (e.g. Cai et al. 2018; Pool et al. 2015; Alvarez-Socorro et al. 2015; Fricke 2016; Luu and Lux 2018, 2019). Briefly, for each pair of nodes in the same set in the bipartite network, the similarity between them is based on the similarity of their neighborhoods. For instance, in the weighted version, we can, respectively, define the similarity matrix of banks and that of firms based on the generalized Jaccard index as

$$S_{ij}^{B-B} = \frac{\sum_{k=1}^{N_F} \min\left(w_{ik}^{bf}, w_{jk}^{bf}\right)}{\sum_{k=1}^{N_F} \max\left(w_{ik}^{bf}, w_{jk}^{bf}\right)}, \quad (\forall 1 \leq i \neq j \leq N), \tag{1}$$

and

$$S_{ij}^{F-F} = \frac{\sum_{k=1}^{N} \min\left(w_{ki}^{bf}, w_{kj}^{bf}\right)}{\sum_{k=1}^{N} \max\left(w_{ki}^{bf}, w_{kj}^{bf}\right)}, \quad (\forall 1 \leq i \neq j \leq N_F). \tag{2}$$

With non-negative weights in the bank-firm credit network, we have that $0 \leq \sum_{k=1}^{N_F} \min(w_{ik}^{bf}, w_{jk}^{bf}) \leq \sum_{k=1}^{N_F} \max(w_{ik}^{bf}, w_{jk}^{bf})$ and $0 \leq \sum_{k=1}^{N} \min(w_{ki}^{bf}, w_{kj}^{bf}) \leq \sum_{k=1}^{N} \max(w_{ki}^{bf}, w_{kj}^{bf})$. Furthermore, since in the present study we focus on banks that lend to at least one firm and firms that borrow from at least one bank, the two denominators $\sum_{k=1}^{N_F} \max(w_{ik}^{bf}, w_{jk}^{bf})$ and $\sum_{k=1}^{N} \max(w_{ki}^{bf}, w_{kj}^{bf})$ are strictly positive. Therefore, S_{ij}^{B-B} and S_{ij}^{F-F} range in [0, 1].

Notice that $S_{ij}^{B-B} = 1$ if and only if $w_{ik}^{bf} = w_{jk}^{bf}$ for all firms k. Furthermore, $S_{ij}^{B-B} = 0$ if and only if the portfolios of the two banks i and j have no overlaps at all. Therefore, each element of the matrix S^{B-B} demonstrates the degree of similarity between the lending portfolios of a pair of banks. In a similar vein, each entry of the matrix S^{F-F} indicates the degree of similarity between the borrowing portfolios of two firms.

Furthermore, the one-mode projection matrices obtained from the original bank-firm credit network also indicate the lending portfolio overlaps between banks or the borrowing portfolio overlaps between firms (e.g. Luu and Lux 2018, 2019). More specifically, defining

$$A^{B-B} = A^{bf} A^{fb}, \tag{3}$$

we then obtain a matrix in which each off-diagonal element is the number of firms

[7] Depending on the considered layer, it can be the long-term loan, the short-term loan, or the total loan.

that a pair of banks co-finance. In a similar way, each off-diagonal element of the A^{F-F} matrix, with

$$A^{F-F} = A^{fb}A^{bf}, \tag{4}$$

indicates the number of common lenders that a pair of firms have.

A major disadvantage of the network projection approach is that the information about the monetary value of loans from banks to firms is partially or completely discarded (e.g. see Lehmann et al. 2008; Luu and Lux 2018). In addition, often sparser information about the original bipartite structure of credit links is projected into denser one-mode networks. More importantly, from an economic perspective, the network-based analysis of the two projection matrices (and the similarity matrices) alone cannot help to identify which pairs of banks tend to be more/less correlated in lending to firms. This provides the motivation for introducing an appropriate procedure used to measure and analyse the loan portfolio correlations between banks, which shall be briefly explained in the following.[8]

2.2.2 Analysis of Portfolio Correlations

2.2.2.1 Correlation Matrix We now introduce the method used to measure the empirical cross-correlation matrix of banks' loan portfolios. Here we focus on the weighted version of bank-firm credit relationships. Additional explanations and results for the binary version (i.e. based on the elements of the matrix $A^{bf} = \{a_{ij}^{bf}\}_{N\text{x}N_F}$) are provided in the "Appendix".

Similar to the procedure often implemented in the analysis of stock returns/prices, for $w_{i,j}^{bf} > 0$, the normalized amount of the loan (in log-scale) from bank i to firm j is defined as

$$\mathrm{X}_{ij} = \frac{\ln\left(w_{ij}^{bf}\right) - \langle \ln\left(w_{ij}^{bf}\right)\rangle}{\sigma_{ij}}, \tag{5}$$

where for each bank i, $\langle \ln(w_{ij}^{bf})\rangle$ and σ_{ij} are the average and the standard deviation of $\ln(w_{ij}^{bf})$ over all N_F firms, respectively. As a special case, we let $\mathrm{X}_{ij} = 0$ if $w_{ij}^{bf} = 0$.[9]

Denote $\mathcal{X} = \{X_{ij}\}_{N\text{x}N_F}$, the cross-correlation matrix of N banks is then defined as

[8] Note that although in the present work we focus on loan portfolio correlations between banks, we also summarize the main results for loan portfolio overlaps and similarities in the "Appendix".

[9] It should be emphasized that in general, almost all main conclusions from this study still hold if we define

$$X_{ij} = \frac{\ln\left(\tilde{w}_{ij}^{bf}\right) - \langle \ln\left(\tilde{w}_{ij}^{bf}\right)\rangle}{\tilde{\sigma}_{ij}},$$

with $\tilde{w}_{ij}^{bf} = (w_{ij}^{bf} + 1)$. One can also use the original weight w_{ij}^{bf} instead of the loan amount in log-scale to define the normalized quantity X_{ij}. Although for the sake of brevity the results for these alternative definitions of X_{ij} are not reported here, they available from the author upon request.

$$C = \{C_{ij}\}_{NxN} = \frac{1}{N_F}\mathcal{X}\mathcal{X}^{\mathcal{T}}, \tag{6}$$

where $\mathcal{X}^{\mathcal{T}}$ is the transposition of $\mathcal{X}$. Note that $-1 \leq C_{ij} \leq 1$ (for $1 \leq i < j \leq N$) and $C_{ii} = 1$ (for all i from 1 to N). From an economic point of view, the value of C_{ij} indicates the correlation between bank i and bank j in lending to all firms. In particular, $C_{ij} > 0$ (< 0) implies that two banks i and j are positively (negatively) correlated, while $C_{ij} = 0$ demonstrates that there is no correlation between the two banks at all.

Note that with the particular application to the bank-firm credit network of Japan in the present work, the ratio $\frac{N_F}{N}$ is larger than 1, since the number of banks is always less than that of firms over the sample period. However, if the number of borrowers was less than the number of lenders, the estimated covariance matrix between banks in lending to firms would become singular.[10] Similarly, in analyses of economic and financial time series, when the number of observations over a relatively short period is smaller than the number of the indices (e.g. those represent different stock prices or macroeconomic indicators), such a singularity problem also emerges. In this case, the estimated covariance matrix will have the singular Wishart distribution (e.g. see Srivastava 2003; Bouchaud et al. 2007; Bodnar et al. 2013, 2014). As a an important consequence in portfolio analysis, if the estimated covariance matrix is singular, the matrix inversion operation used to estimate the portfolio weights is not possible. To tackle this problem, several transformation methods with practical applications to portfolio theory haven been proposed (c.f. Pappas and Kaimakamis 2010; Bodnar et al. 2016, 2017, 2018, 2019, among others for a more detailed discussion on different transformation methods and their practical relevance).

2.2.2.2 Random Matrix Theory and Selection of Factors We shall now introduce the relevant results of RMT that will be used to analyse the cross-correlations matrix defined in Eq. (6). Suppose that $\{\lambda_i\}_{i=1}^{i=N}$ are the eigenvalues of a correlation matrix C and $f_C(\lambda)$ is the probability density function of the eigenvalues of C. According to RMT (e.g. Marčenko and Pastur 1967; Tao and Vu 2012), if the elements of $\mathcal{X} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$, then in the limit $N, N_F \to \infty$ with $Q = \frac{N_F}{N} \to a$ (a is constant and $a > 1$), the eigenvalue distribution $f_C(\lambda)$ is given by the Marchenko-Pastur law

$$f_C(\lambda) = \begin{cases} \frac{Q}{2\pi\sigma^2}\frac{\sqrt{(\lambda^*_{max} - \lambda)(\lambda - \lambda^*_{min})}}{\lambda} & \text{if } \lambda^*_{min} \leq \lambda \leq \lambda^*_{max}, \\ 0 & \text{elsewhere.} \end{cases} \tag{7}$$

In (7), σ^2 is the variance of the elements of $\mathcal{X}$ (σ^2 is equal to 1 in our case), and λ^*_{max} and λ^*_{min} are the upper and lower bounds of eigenvalues predicted by RMT, respectively. With unit variance, these two bounds are given by

[10] This could also happen in other bipartite financial networks such as bank-asset or investor-asset networks. In particular, when the number of assets is smaller than the number of banks or investors, it also leads to the singularity problem in the bank-bank covariance matrix or the investor-investor covariance matrix.

$$\lambda^*_{max} = (1 + \sqrt{1/Q})^2 \text{ and } \lambda^*_{min} = (1 - \sqrt{1/Q})^2. \tag{8}$$

Generally speaking, if some of the eigenvalues of the empirical cross-correlation matrix are beyond the interval $[\lambda^*_{min}, \lambda^*_{max}]$, they carry genuine information about the correlations between banks (Plerou et al. 2000, 2002). Nevertheless, it should be emphasized that the presence of extreme values in the elements of $\mathcal{X}$ (e.g. when $\mathcal{X}$ is a heavy-tailed random matrix) might cause the largest eigenvalues to exceed the upper bound implied by the Marchenko-Pastur law (e.g. Biroli et al. 2007; Bouchaud and Potters 2009). In this case, some eigenvalues that are larger than λ^*_{max} might actually contain no genuine information.

As shown in Biroli et al. (2007), whenever the largest entry of $\mathcal{X}$ (in absolute terms), S, satisfies that $S \leq (N.N_F)^{\frac{1}{4}}$, the upper limit for the largest eigenvalue still remains stuck at λ^*_{max}. However, if $S > (N.N_F)^{\frac{1}{4}}$, a new upper bound $\lambda^{(1)}_{max}$ will emerge:

$$\lambda^{(1)}_{max} = \left(\frac{1}{Q} + \frac{S^2}{N_F}\right)\left(1 + \frac{N_F}{S^2}\right). \tag{9}$$

Furthermore, when the elements of $\mathcal{X}$ have power-law tails with exponent μ, then the large values among $\{X_{ij}\}$ will dominate the top largest eigenvalues. The new upper bound $\lambda^{(2)}_{max}$ now depends on μ as

$$\lambda^{(2)}_{max} = N^{4/\mu-1} Q^{2/\mu-1}. \tag{10}$$

Hence, in the present study, to capture the potential effects of extreme values of loans from banks to firms, we also compare the top largest eigenvalues with these different bounds. As shown later, in our analysis the top five largest eigenvalues of the empirical correlation matrix C are always larger than all three upper bounds λ^*_{max}, $\lambda^{(1)}_{max}$ and $\lambda^{(2)}_{max}$.

2.2.2.3 Decomposition of Correlation Matrix We now introduce one of the most important applications of RMT, i.e. how to extract latent information from the largest eigenvalues deviating from the random bulk. Without loss of generality, we assume that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ are the eigenvalues of the empirical correlation matrix C, and $u_1, u_2, \ldots, u_N$ are their corresponding eigenvectors. The correlation matrix C can be diagonalized as

$$C = \mathrm{U}\Lambda\mathrm{U}^{\mathcal{T}}, \tag{11}$$

with $\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_N\}$, and $\{\mathrm{U}\}_{NxN}$ is an orthonormal matrix, whose i^{th} column is the normalized eigenvector u_i corresponding to the eigenvalue λ_i. As a result, for each λ_i we have

$$\lambda_i = \mathrm{u}_i^{\mathcal{T}} C \mathrm{u}_i = \mathrm{u}_i^{\mathcal{T}} \mathrm{Cov}(\mathcal{X})\mathrm{u}_i = \mathrm{Var}(\mathrm{u}_i^{\mathcal{T}} \mathcal{X}). \tag{12}$$

Let $y_i = \mathrm{u}_i^{\mathcal{T}} \mathcal{X}$, we obtain the following property for the total variance of $\{\mathcal{X}_i\}_{i=1}^N$:

$$\sum_{i=1}^{N} \text{Var}(\mathcal{X}_i) = N = \sum_{i=1}^{N} \lambda_i = \sum_{i=1}^{N} \text{Var}(y_i). \tag{13}$$

From Eqs. (12) and (13), we can easily show that the ratio $\frac{\lambda_i}{N}$ indicates the percentage of the total variance $\sum_{i=1}^{N} \text{Var}(\mathcal{X}_i)$ explained by the principal component $y_i = \mathrm{u}_i^{\mathcal{T}} \mathcal{X}$ (Jolliffe 1986; Wang et al. 2011; Gewers et al. 2018).

In addition, since Eq. (11) can be re-written as

$$C = \sum_{i=1}^{N} \lambda_i u_i u_i^{\mathcal{T}}, \tag{14}$$

we shall now decompose the empirical correlation C into different parts capturing the effects of different factors. First, if the largest eigenvalue λ_1 deviates significantly from the random bulk, the so-called "market" mode, i.e. the global common factor (or the systematic factor) that influences the lending of all banks in the credit market, can be extracted from this eigenvalue and its corresponding eigenvector (Plerou et al. 2000, 2002). Its effect on C is represented by

$$C^m = \lambda_1 u_1 u_1^{\mathcal{T}}, \tag{15}$$

and the rest of C is

$$\text{FC}^m = C - C^m, \tag{16}$$

which is the filtered matrix after the market-wide effect is subtracted.

Next, genuine information can be extracted from the next largest eigenvalues if they are still larger than the selected upper bound (e.g. the bound λ_{max}^{*} implied by RMT or the more "conservative" one given by $\lambda_{max}^{all} = \max(\lambda_{max}^{*}, \lambda_{max}^{(1)}, \lambda_{max}^{(2)})$). Such information may be associated with the level of sub-groups (namely the "groups" modes) (e.g. Kim and Jeong 2005; Shen and Zheng 2009; Jiang and Zheng 2012). In that case, following Kim and Jeong (2005), we can further decompose C into three parts C^m, C^g, and C^r as

$$C = C^m + C^g + C^r, \tag{17}$$

where C^g and C^r indicate the effect of the "groups" modes and the rest of C, respectively. From an economic perspective, the matrix C^g captures the influence of non-systematic components (e.g. sectoral factors, regional factors, etc.) that only affect the lending of a subset of banks in the credit market. The residual represented by the matrix C^r is refereed to as the random noise part.[11]

[11] Indeed, the interpretations for different factors used in the decomposition of the matrix C are somewhat similar to those used for the dynamic factors models, which have also become one of the important dimension reduction techniques in macro-econometrics as well as financial econometrics (e.g. Forni et al. 2000; Forni and Lippi 2001; Barigozzi et al. 2014; Forni et al. 2015; Matteo and Marc 2016b, a; Forni et al. 2017).

Mathematically,

$$C^g = \sum_{i=2}^{N_g} \lambda_i u_i u_i^{\mathcal{T}}, \tag{18}$$

and the rest

$$C^r = \sum_{i=N_g+1}^{N} \lambda_i u_i u_i^{\mathcal{T}}, \tag{19}$$

where N_g is determined based on the number of the next largest eigenvalues larger than the selected upper bound (see also Kim and Jeong (2005), MacMahon and Garlaschelli (2015) for further discussions about the selection of N_g). From Eqs. (17) to (19), we can define $\mathrm{FC}^{m,g} = (C - C^m - C^g) = C^r$ as the filtered matrix, in which the effects of the "market" and "groups" modes are subtracted.

2.2.2.4 Absorption Ratios and Systemic Risk From Eq. (13), for each $i = 1, 2, \ldots N$, we define the absorption ratio E_i as

$$E_i = \frac{\sum_{j=1}^{i} \lambda_j}{N} = \frac{\sum_{j=1}^{i} \mathrm{Var}(\mathcal{X}_j)}{N}, \tag{20}$$

which indicates the fraction of the total variance of $\{\mathcal{X}_j\}_{j=1}^{N}$ explained by (or absorbed by) the first i principal components (e.g. Pukthuanthong and Roll 2009; Wang et al. 2011; Billio et al. 2012).

Since $\sum_{i=1}^{N} \mathrm{Var}(\mathcal{X}_i) = N = \sum_{i=1}^{N} \lambda_i$, the ratio E_N is equal to 1, representing 100% of the total variance explained by all components together. Hence, comparing different absorption ratios $E_1, \ldots, E_k$ to $E_N = 1$ allows us to assess the contribution of the significant principal components to the total variance.[12] On top of this, we can use the absorption ratios to analyse the systemic risk in the market. For example, if E_1 accounts for a larger part of E_N when the largest eigenvalue λ_1 increases, it indicates a higher level of systemic risk. This is because it implies that banks in the credit market now become more unified and strongly coupled, and a shock, thus, can be quickly and widely propagated throughout the whole system (e.g. Pukthuanthong and Roll 2009; Billio et al. 2012; Zheng et al. 2012; Meng et al. 2014).

2.2.2.5 Inverse Participation Ratios We use the Inverse Participation Ratio (IPR) as an indicator of the contribution to the eigenvectors of the correlation matrix (Plerou et al. 1999, 2002). It is given by

$$\mathrm{IPR}_i = \sum_{j=1}^{N} u_i(j)^4. \tag{21}$$

The inverse of IPR_i indicates the number of elements that contribute significantly to

[12] Here k is the largest integer such that λ_k is larger than the selected bound, e.g. λ^*_{max} implied by RMT or the more "conservative" one $\lambda^{all}_{max} = \max(\lambda^*_{max}, \lambda^{(1)}_{max}, \lambda^{(2)}_{max})$.

u_i. More specifically, a lower value for IPR implies that banks contribute more equally (i.e. a less localized behavior). In contrast, a higher value for IPR shows that a smaller number of banks dominate in the eigenvector (i.e. a more localized behavior), signaling the presence of a more hierarchical structure.

In the normalized version of eigenvectors, if $\sum_{j=1}^{N} u_i(j)^2 = 1$ $(\forall i = 1, 2, \ldots N)$, the ratio IPR_i will have two limiting cases

$$1 \geq \text{IPR}_i \geq \frac{1}{N}, \ \forall i = 1, 2, \ldots N. \tag{22}$$

Notice that $\text{IPR}_i = \frac{1}{N}$ if and only if $u_i(j) = \sqrt{\frac{1}{N}}$ for all $j = 1, 2, \ldots N$ (i.e. under the completely homogeneous structure). In contrast, $\text{IPR}_i = 1$ if and only if only one element contributes to the eigenvector u_i while all other elements are equal to zero (i.e. under the most hierarchical structure).

2.2.2.6 Complex Network Approach to Correlation Structure Analysis In order to identify the backbone of the interdependencies between banks, we use the so-called Planar Maximally Filtered Graph (PMFG). Broadly speaking, this method allows us to filter and map the correlation matrix into a sparser graph and retain only the important correlation parts (e.g. Tumminello et al. 2005; Di Matteo et al. 2010; Aste et al. 2010; Jiang et al. 2014; Massara et al. 2015). Based on the decomposition of the empirical correlations into different eigenmodes (see Eq. (17)), we can also construct the PMFG for the different correlation matrices, e.g. the filtered correlation matrix preserving only the "market" mode (i.e. C^m), the filtered correlation matrix retaining only the "groups" modes (i.e. C^g), or the noise-filtered correlation matrix maintaining all significant correlations (i.e. $C^{m,g} = C^m + C^g = C - C^r$).

Indeed, the construction algorithm of the PMFG is similar to the one used for the so-called Minimal Spanning Tree (MST)–one of the simplest connected graphs widely used to capture a minimal set of relevant interactions associated with the strongest correlations (e.g. Mantegna 1999). To construct the MST graph, one can follow the following steps: First, sorting the pairwise correlations $\{C_{ij}\}$ in descending order, we obtain a list $\mathcal{L}$ of the N^2 reordered correlation elements. In the second step, choosing the two banks i and j associated with the first element in $\mathcal{L}$, we put an edge between banks i and j to the MST graph. In the third step, we select the next edge in the list $\mathcal{L}$, and then add the corresponding link and nodes to the graph if it is not connected to the existing nodes and does not form any cycles. We then repeat the third step until all the correlation elements in the list $\mathcal{L}$ have been examined.

Notice that since in the third step in the construction of the MST, all new edges are excluded if they form a cyclic structure, the resulting filtered MST is a tree graph consisting of N banks with $(N - 1)$ edges. Under such a strong constraint on an acyclic structure, the presence of significant correlations that have loops will be definitely discarded (e.g. Tumminello et al. 2005; Di Matteo et al. 2010; Wang et al. 2018). To provide a more significant and richer structure, Tumminello et al.

(2005) propose the PMFG in which the resulting connected planar graph retains $(3N-6)$ edges and possibly contains 3- and 4- closed cliques (in the third step of the aforementioned construction algorithm of the MST). In fact, it can be shown that the resulting MST graph is always a subgraph of the PMFG (e.g. Tumminello et al. 2005; Pozzi et al. 2008).

3 Portfolio Correlations in the Bank-Firm Credit Market of Japan

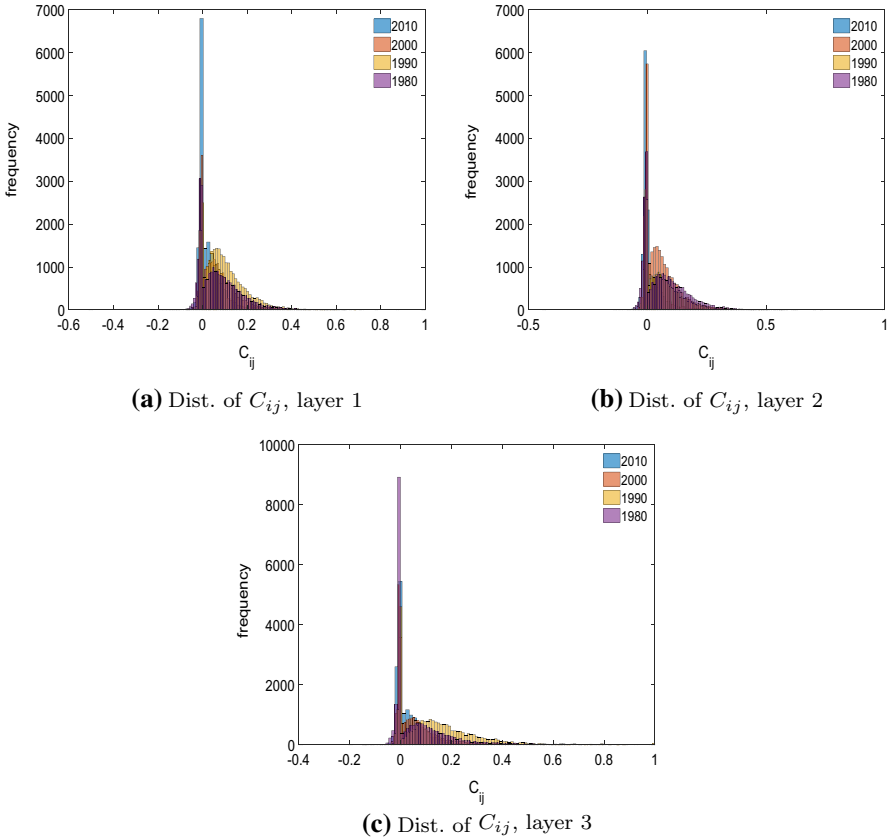
To begin with, we briefly summarize the main statistical properties of the elements of the correlation matrices in the three lending layers of the bank-firm credit market of Japan over the period from 1980 to 2012. The time-varying distributions of C_{ij} in the different layers are shown in Fig. 1, while the fundamental statistics including the mean, median, skewness and kurtosis of C_{ij} are reported in Fig. 2. Overall, we observe that the averages of correlations in all layers are relatively small but always positive. In addition, these distributions exhibit positive skewness and high kurtosis.

In order to investigate whether non-random patterns can be identified in the correlations between loan portfolios, in each layer, first, we compute the eigenvalues as well as eigenvectors of the empirical correlation matrix C. After that, we will compare the distribution of eigenvalues of C with the one predicted by RMT. Note that in our analysis we also consider the potential effects of extreme values on the upper limit of eigenvalues by taking into account the other two upper bounds $\lambda_{max}^{(1)}$ and $\lambda_{max}^{(2)}$, besides the bound λ_{max}^{*} implied by RMT.

As shown in Fig. 3 and in Fig. 11 (see the "Appendix"), a majority of the eigenvalues of the empirical correlation matrix lie within the random bulk $[\lambda_{min}^{*}, \lambda_{max}^{*}]$. Compared to this, we find that about additional 10% of the eigenvalues are in the broader range $[\lambda_{min}^{*}, \lambda_{max}^{all}]$ when the more conservative upper bound $\lambda_{max}^{all} = \max(\lambda_{max}^{*}, \lambda_{max}^{(1)}, \lambda_{max}^{(2)})$ is used. Furthermore, we also observe a group of the smallest eigenvalues deviating from the RMT's left edge (i.e. less than the lower bound λ_{min}^{*}). However, these eigenvalues are not of interest in the present work as their contribution to the total variance is relatively small.

Moving on to the largest eigenvalues exceeding the RMT's right edge, typically from 10 % to 20 % of eigenvalues are always larger than λ_{max}^{*} over the years, while a smaller fraction is observed when we only count the eigenvalues that exceed λ_{max}^{all}. All taken together, the inspection of the empirical distribution the eigenvalues suggests that genuine information from the bank-bank correlation matrices and significant factors driving the correlation dynamics over the years in the three layers can be extracted from some of the largest eigenvalues.

In the following, for the sake of conciseness, we select the top five largest eigenvalues of the empirical correlation matrix in each lending layer and then investigate their temporal evolution over time. As shown in the panels (a), (c), and (e) of Fig. 4, over the years, these eigenvalues persistently deviate from the random bulk, and they are always larger than the all three upper bounds λ_{max}^{*}, $\lambda_{max}^{(1)}$ and $\lambda_{max}^{(2)}$. To show the contribution of the first five principal components to the total variance and to infer the systemic risk, we also report the associated absorption ratios defined



(a) Dist. of C_{ij}, layer 1

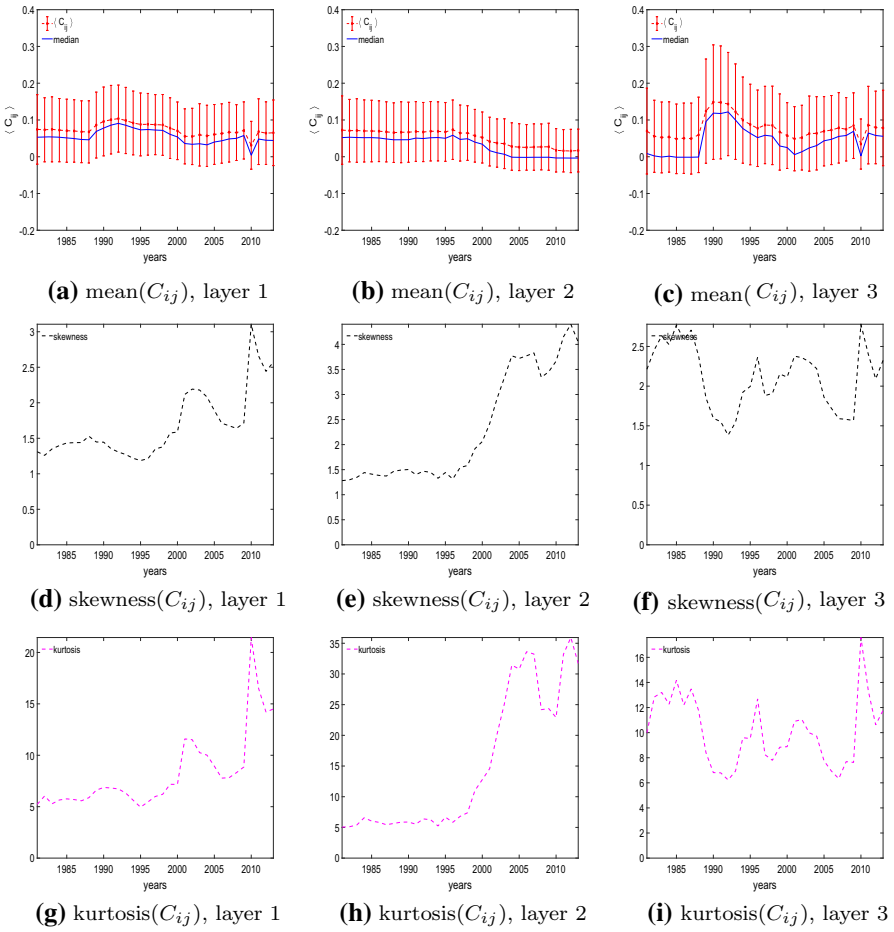
(b) Dist. of C_{ij}, layer 2

(c) Dist. of C_{ij}, layer 3

Fig. 1 Temporal evolution of the distributions of C_{ij} in the three layers. Panels **a–c**: The distributions of C_{ij} in the three layers in the four selected years 1980, 1990, 2000, 2010

in Eq. (20). Overall, to a certain degree, the layer 1 (total loans) and layer 3 (long-term loans) display a similar trend. As demonstrated in Fig. 4, a couple of regimes in the bank-firm credit market of Japan can be identified from these two layers. For example, the temporal evolution shows that almost all top five eigenvalues and the absorption ratios increase sharply during the period of the lending boom and Japanese asset price bubble (1986–1991).[13] This implies that in this period the top five significant factors become more important driving forces of the correlation dynamics, and Japanese banks tend to be more tightly coupled in lending to the non-financial sector of the economy. Hence, it indicates a higher level of systemic risk in the credit market during the bubble time in the sense that banks' lending becomes more vulnerable to negative shocks (Kritzman et al. 2011; Billio et al. 2012; Meng et al. 2014). As we can see in the consecutive years after the bubble bursting, all of these eigenvalues as well as the corresponding absorption ratios substantially

[13] See, for example, Hoshi and Kashyap (2000), Posen (2003) and Brunnermeier and Schnabel (2016) for further discussions on the lending boom and the emergence of asset price bubble in Japan in 1980s.



(a) mean(C_{ij}), layer 1 **(b)** mean(C_{ij}), layer 2 **(c)** mean(C_{ij}), layer 3

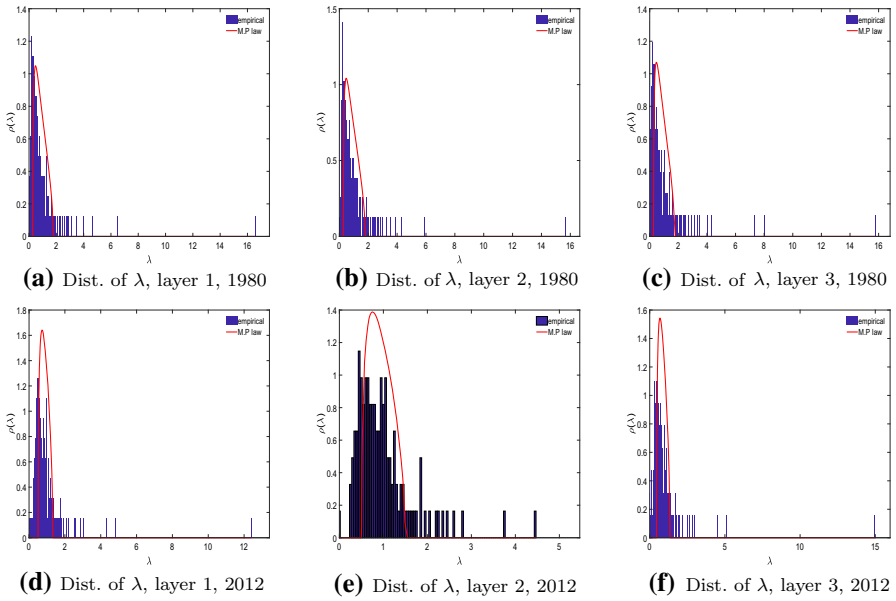
(d) skewness(C_{ij}), layer 1 **(e)** skewness(C_{ij}), layer 2 **(f)** skewness(C_{ij}), layer 3

(g) kurtosis(C_{ij}), layer 1 **(h)** kurtosis(C_{ij}), layer 2 **(i)** kurtosis(C_{ij}), layer 3

Fig. 2 Fundamental statistics of C_{ij} in the three layers. Panels **a–c** The averages (together with their 95 % confidence intervals) and the medians of C_{ij} in the three layers. Panels **d–f** The values of the skewness of C_{ij} in the three layers. Panels **g–i** The values of the kurtosis of C_{ij} in the three layers

decline, and even decrease further after the Asian financial crisis (1997–1998). Such a trend somewhat stops in 2002, when the Bank of Japan (BOJ) implements a monetary stimulus policy (e.g. Bowman et al. 2015). In addition, as a consequence of the economic contraction in Japan hit by the recent global crisis (e.g. Kawai and Takagi 2011), we observe a significant decrease in all of these eigenvalues and absorption ratios in 2009 in all three layers.

Furthermore, interestingly, compared to the layers 1 and 3, the layer 2 (short-term loans) possesses a different dynamics: the largest eigenvalues and their contributions to the total variance show a decline from the Asian financial crisis, suggesting that from then on Japanese banks tend to be less unified in this layer. Notice that as shown in panel (b) of Fig. 2, the correlations between banks in the layer 2 also exhibit a similar dynamics: on average, banks' short-term loans to firms become less



(a) Dist. of λ, layer 1, 1980 **(b)** Dist. of λ, layer 2, 1980 **(c)** Dist. of λ, layer 3, 1980

(d) Dist. of λ, layer 1, 2012 **(e)** Dist. of λ, layer 2, 2012 **(f)** Dist. of λ, layer 3, 2012

Fig. 3 Distributions of the eigenvalues of the correlation matrices, compared with RMT, in the three layers, in 1980 and 2012. Panels **a**–**c** show the distributions of of λ in 1980, and **d**–**f** show the distributions of λ in 2012. In all panels, the red curve represents the spectral distribution explained by the Marchenko-Pastur law. Note that over the years, we observe that the top five largest eigenvalues of the empirical correlation matrix in each lending layer are always larger than the three upper bounds λ^*_{max}, $\lambda^{(1)}_{max}$ and $\lambda^{(2)}_{max}$ (see also Fig. 11 in the "Appendix")

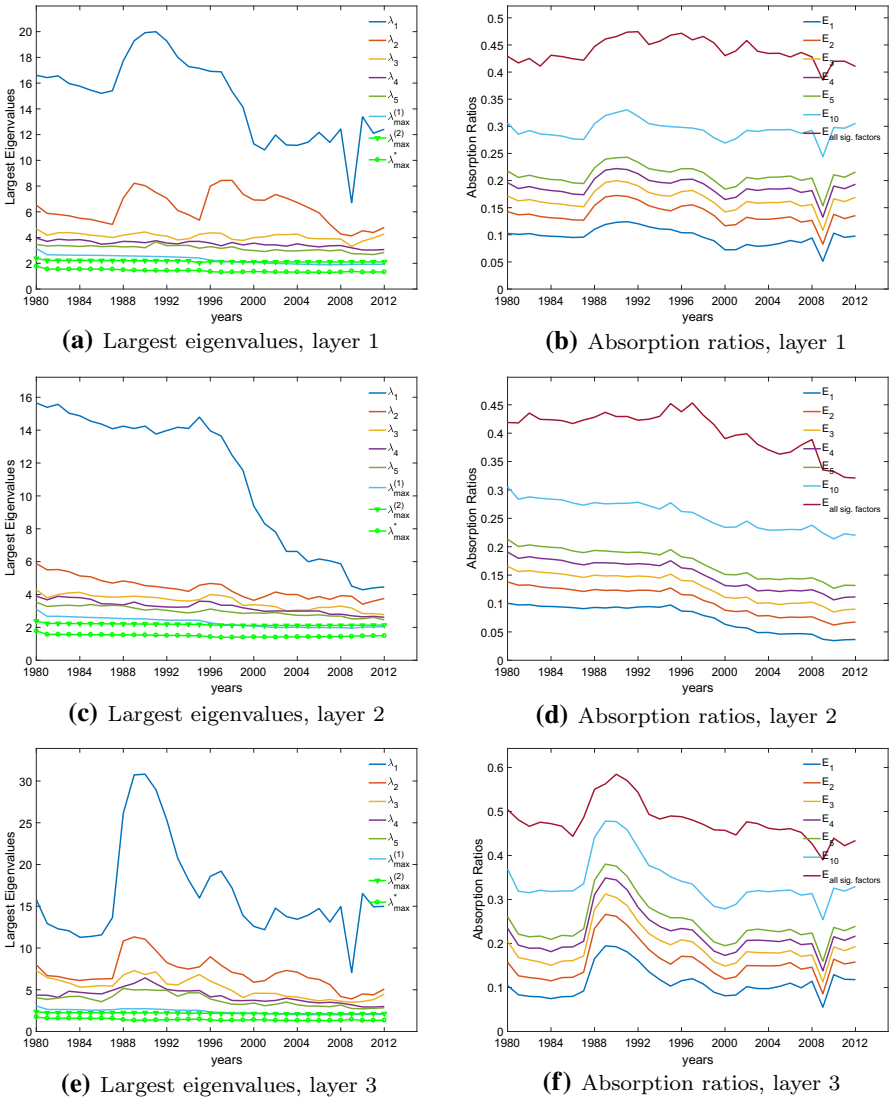
correlated since the aftermath of that crisis. In fact, by a closer inspection from a network perspective, we find that credit linkages in this lending layer also become sparser and the average of weights (in absolute terms and in log-scale) substantially drops since then.

Having discussed the important statistics of the eigenvalues of the correlation matrix C in each lending layer, we now move on to the analysis of the eigenvectors of C. Note that the i^{th} element of an eigenvector u_k indicates the contribution of the bank i to u_k, thus, showing the role of the bank i in the k^{th} principle component.

To begin with, we compare the distributions of the empirical eigenvector elements with the theoretical distribution predicted by RMT. According to the Gaussian prediction of RMT, the elements of each normalized eigenvector will follow (e.g. Guhr et al. 1998; Laloux et al. 1999; Plerou et al. 2002):

$$P^{RMT}(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}. \tag{23}$$

The detailed comparison results are shown in Fig. 12 in the "Appendix". Overall, again we typically find that the distributions of the elements of eigenvectors associated with the largest eigenvalues significantly deviate from the Gaussian prediction of RMT. In contrast, the distributions of the elements of eigenvectors associated with eigenvalues in the random bulk are somewhat more similar to P^{RMT}. This



(a) Largest eigenvalues, layer 1

(b) Absorption ratios, layer 1

(c) Largest eigenvalues, layer 2

(d) Absorption ratios, layer 2

(e) Largest eigenvalues, layer 3

(f) Absorption ratios, layer 3

Fig. 4 Evolution of the top five largest eigenvalues and absorption ratios for the correlation matrices in the three layers. All panels on the left show the evolution of the top five largest eigenvalues. These eigenvalues are persistently larger than the bound λ^{*}_{max} implied by RMT as well as the two more conservative bounds $\lambda^{(1)}_{max}$ and $\lambda^{(2)}_{max}$. The right panels show the five associated absorption ratios (E_1 to E_5) and the absorption ratio of the top ten largest eigenvalues (E_{10}) and that of all eigenvalues that exceed λ^{*}_{max} (namely $E_{\text{allsig.factors}}$)

observation is in agreement with those reported in analyses of stock returns (e.g. Laloux et al. 1999; Plerou et al. 1999).

In the next step, based on the eigenvector elements, we investigate the degree of homogeneity/heterogeneity of banks' contribution to C. To this end, we compute the

Inverse Participation Ratios $\{\text{IPR}_k\}_{k=1}^{k=N}$ and then examine the relationship between these ratios and the corresponding eigenvalues. Large values of IPR_k reveal that only few banks contribute to u_k, suggesting the presence of a localized behavior of banks. In contrast, a small value of IPR_k indicates that many banks contribute to u_k together, showing that the factor (if it exists) extracted from that eigenvector has a pervasive effect in the market.

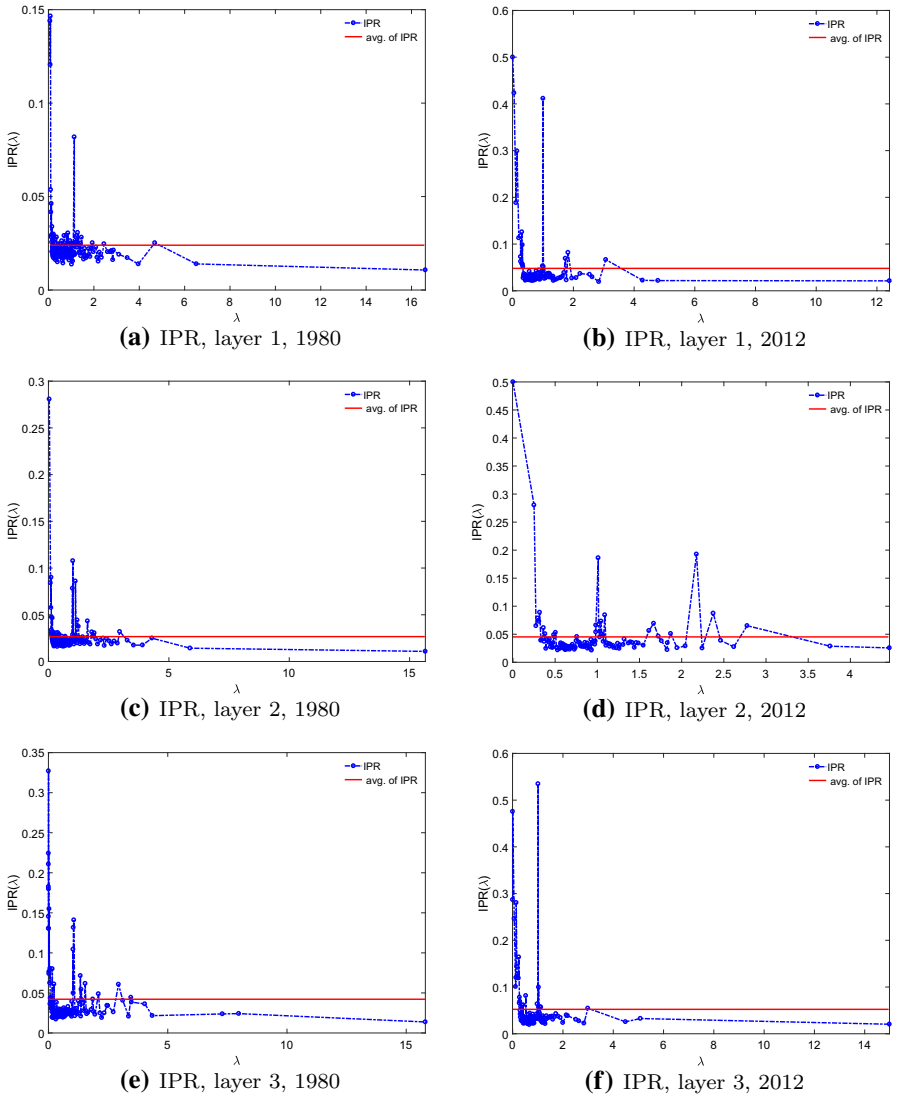
All panels of Fig. 5 demonstrate a common feature among the three layers: some smallest eigenvalues have the highest level of IPR and their IPR ratios deviate from the average of $\{\text{IPR}_k\}_{k=1}^{k=N}$. This indicates that their associated eigenvectors are more localized, i.e. only a small group of banks contribute to them. Nevertheless, these smallest eigenvalues are less than 1 and the contribution of their associated factors/components to the total variance (recalling also Eqs. (12) and (13)) is actually negligible. Moving on to the IPRs associated with the eigenvalues located in the center of the eigenvalue distribution, we find that these eigenvalues often have a relatively low level of IPR. In fact, these features of the smallest eigenvalues and those located at the center of the distribution are analogous to what universally found in similar analyses of stock return correlations (e.g. Plerou et al. 1999).

However, interestingly, we do observe a distinct feature for the eigenvector elements of the largest eigenvalue: over the years, we find that IPR_1 is typically less than the average of $\{\text{IPR}_k\}_{k=1}^{k=N}$, showing that there are many banks contributing to the eigenvector associated with the first largest eigenvalue. From an economic perspective, this indicates a wide effect that the "market" mode has on the banking system. Such a behavior still emerges in some of the next largest eigenvalues deviating from the random bulk. This result implies that besides the global common factor, some "groups" factors also play a certain role in explaining the correlation dynamics. Furthermore, it also signals the presence of large communities (clusters) of banks in which the members within each community (cluster) tend to be more correlated in their lending activities, since they are influenced by some common pervasive factors.

It should be emphasized that under a more detailed inspection, in each year, we also observe a more localized behavior in at least one of the eigenvectors associated with the largest eigenvalues other than λ_1. In this case, the associated factor only drives the co-movements among a smaller subset of banks. For instance, as shown in panels (b), (f) of Fig. 5, in 2012, in the first and the third layers, IPR_4 is relatively larger than the average of $\{\text{IPR}_k\}_{k=1}^{k=N}$, while in the second layer, such a behavior can be found in the case of IPR_8.[14] Nevertheless, since we typically observe distinct localized eigenvectors in different years, it is hard to track the economic imprints responsible for the emergence of this behavior. We therefore leave this issue for future research.

Furthermore, if the largest eigenvalue and the next largest ones stand for the "market" and "groups" modes (i.e. systematic and "non-systematic" factors), respectively, it would be interesting to examine their effects by comparing the raw

[14] In Fig. 13 in the "Appendix", we also show the visualization for a more localized eigenvector and a more homogenized eigenvector as examples.



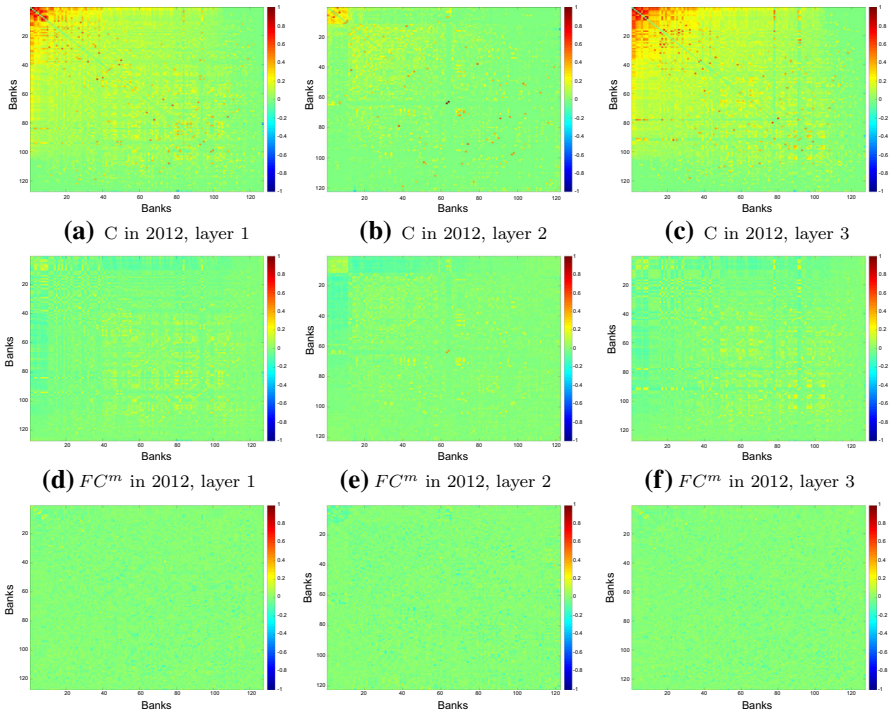
(a) IPR, layer 1, 1980 **(b)** IPR, layer 1, 2012

(c) IPR, layer 2, 1980 **(d)** IPR, layer 2, 2012

(e) IPR, layer 3, 1980 **(f)** IPR, layer 3, 2012

Fig. 5 Inverse Participation Ratios of the eigenvectors of the correlation matrices in the three layers. Panels **a**, **b** show IPR versus λ in the layer 1, in 1980 and 2012. Panels **c**, **d** show IPR versus λ in the layer 2, in 1980 and 2012. Panels **e**, **f** show IPR versus λ in the layer 3, in 1980 and 2012. The red lines stand for the averages of $\{\mathrm{IPR}_k\}_{k=1}^{k=N}$

correlation matrix with the filtered ones. As can be seen in the panels (a) to (f) of Fig. 6, after excluding the influence of the "market" mode on cross-correlation matrix C, some significant correlations still remain in the filtered matrix FC^m. Such significant correlations are actually mainly driven by the "groups" modes associated with the next largest eigenvalues. A further decomposition can be implemented to split the "market" mode-filtered matrix into the correlations driven



(a) C in 2012, layer 1 **(b)** C in 2012, layer 2 **(c)** C in 2012, layer 3

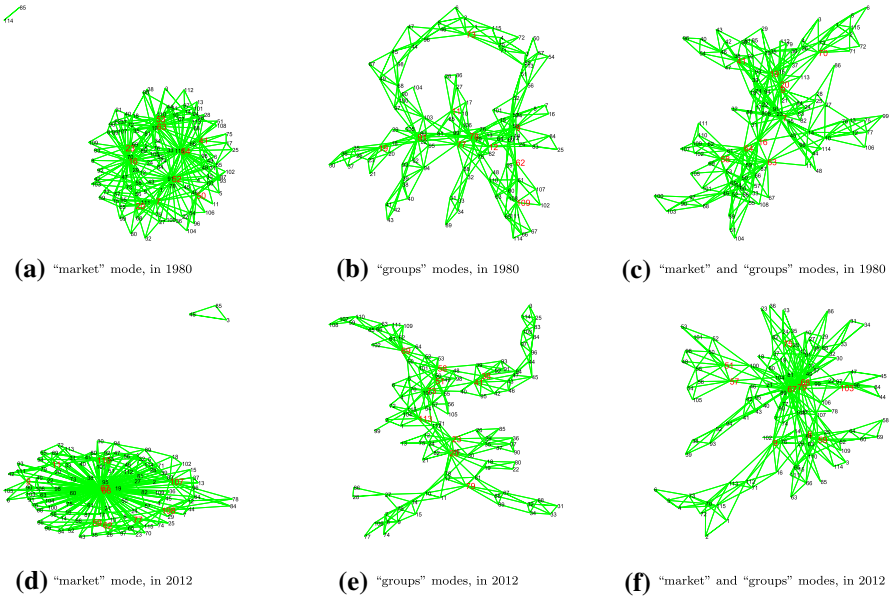
(d) FC^m in 2012, layer 1 **(e)** FC^m in 2012, layer 2 **(f)** FC^m in 2012, layer 3

(g) $FC^{m,g}$ in 2012, layer 1 **(h)** $FC^{m,g}$ in 2012, layer 2 **(i)** $FC^{m,g}$ in 2012, layer 3

Fig. 6 Correlations and filtered correlations between banks in lending to firms in the three layers, in 2012 as an example. The three upper panels from **a** to **c** show the raw correlation matrices. The three panels from **d** to **f** are the filtered correlation matrices, subtracting the influence of the "market" mode. The three panels from **g** to **i** show the filtered correlation matrices, subtracting the effects of the "market" mode as well as "groups" modes. In each layer, banks are sorted in descending order of eigenvector centrality of the raw correlation matrix

by the "groups" modes and by noise. The "market" as well as "groups" modes-filtered matrices $FC^{m,g}$ in three layers are shown in the panels (g) to (h) of Fig. 6. We observe that almost all correlations are significantly subtracted in $FC^{m,g}$.

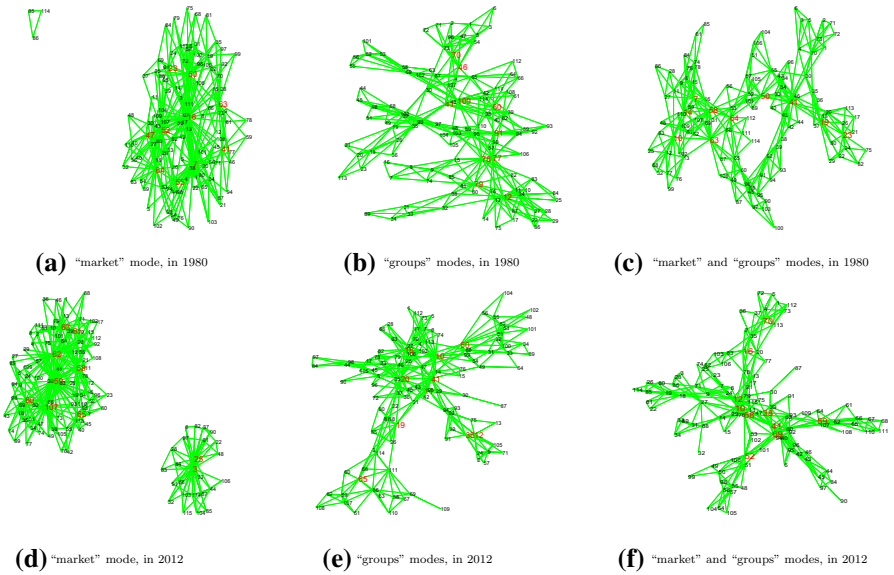
All taken together, the results we have obtained so far demonstrate that the analysis based on the methods of RMT and PCA can be used to identify important, significant patterns in the correlations between banks in lending to firms in the other sectors of the economy. In particular, we have shown that the correlations are mainly driven by a global common factor influencing the lending of all banks and some non-systematic factors that only affect the lending of a subset of banks in the credit market of Japan. To gain a deeper understanding of the structure and the dynamics of these correlations, we make one more step by combining these methods with those of complex networks. More specifically, for each year, we use the PMFG graph to extract the skeleton structure from the different parts of the cross-correlation matrices, i.e. the part represents the effects of "market" mode (C^m), the part captures the effects of "groups" modes (C^g), the part represents the effects of



(a) "market" mode, in 1980 **(b)** "groups" modes, in 1980 **(c)** "market" and "groups" modes, in 1980

(d) "market" mode, in 2012 **(e)** "groups" modes, in 2012 **(f)** "market" and "groups" modes, in 2012

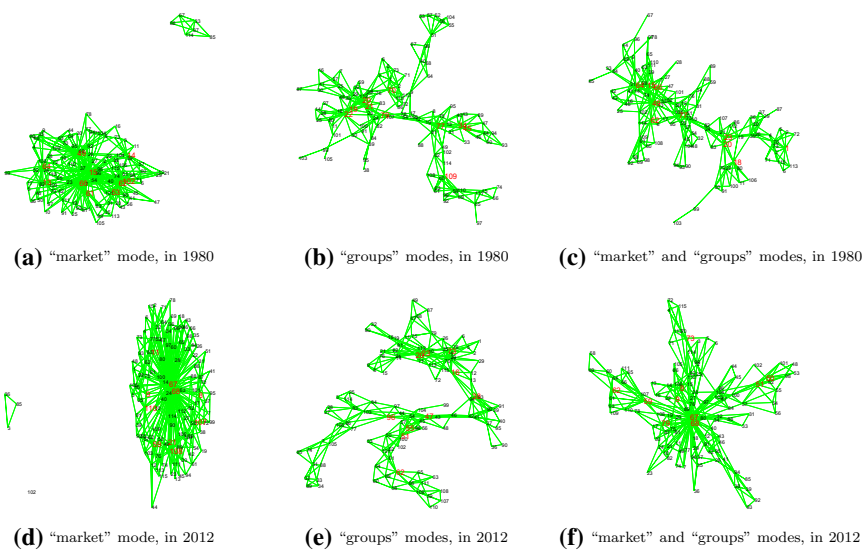
Fig. 7 The interaction structure between banks for C^m, C^g, and $C^{m,g}$ obtained with the PMFG graph, in layer 1



(a) "market" mode, in 1980 **(b)** "groups" modes, in 1980 **(c)** "market" and "groups" modes, in 1980

(d) "market" mode, in 2012 **(e)** "groups" modes, in 2012 **(f)** "market" and "groups" modes, in 2012

Fig. 8 The interaction structure between banks for C^m, C^g, and $C^{m,g}$ obtained with the PMFG graph, in layer 2



(a) "market" mode, in 1980 **(b)** "groups" modes, in 1980 **(c)** "market" and "groups" modes, in 1980

(d) "market" mode, in 2012 **(e)** "groups" modes, in 2012 **(f)** "market" and "groups" modes, in 2012

Fig. 9 The interaction structure between banks for C^m, C^g, and $C^{m,g}$ obtained with the PMFG graph, in layer 3

both "market" as well as "groups" modes ($C^{m,g} = C^m + C^g$). We then study the persistence and changes in the skeleton structure over the years.

For the illustration purpose, in Figs. 7, 8 and 9, we show PMFG graphs extracted from different parts of the correlation matrices in the two years 1980 and 2012. In each panel of these Figures, each number corresponds to a bank, and each link (in green color) indicates the presence of a significant relation between a pair of banks. The numbers in red color represents the hub banks that have the highest levels of the degrees in the PMFG graph. Overall, the correlations between banks reveal different backbone structures in different periods. Indeed, by having a closer inspection over the years (for example, see Figs. 17 and 18 in the "Appendix") we often observe that when the dominance of a group of banks in one period gradually disappears, the credit market starts to build-up a different structure in the next period in which another group of banks emerge as the key players in the backbone of the cross-correlations.

4 Concluding Remarks

In this paper, we have examined the structure and the dynamics of the correlation matrices for the banks' loan portfolios in a large data set of the bank-firm credit network of Japan during the period from 1980 to 2012. Our results show that only a subset of the eigenvalues contain genuine information while the rest actually corresponds to noise or just have a small effect on the correlations between banks in lending to firms in the other sectors of the economy. This implies that the methods of RMT and PCA can be used for filtering noise from the empirical correlation matrices of banks' loan portfolios.

The dynamics of the largest eigenvalues and the associated absorption ratios reveal different milestones in the credit market of Japan over the years. More specifically, there is a dramatic increase in the level of systemic risk during the period of the Japanese asset price bubble (1986–1991). In contrast, we observe a significant decrease in all of these eigenvalues and the associated absorption ratios in 2009, which might be considered as a consequence of the economic contraction in Japan hit by the global crisis (e.g. Kawai and Takagi 2011).

Based on the eigenvector elements, we have further investigated the localization and clustering behaviours of banks. One the one hand, we find that in the case of the eigenvectors corresponding to a group of smallest eigenvalues, there are only few banks contributing to them. Such a localization behavior is similar to what is often found in analyses of stock price changes (e.g. Plerou et al. 1999). However, since these eigenvalues are smaller than 1 and the top largest eigenvalues, the contribution of their corresponding principal components to the total variance is actually negligible. On the other hand, we find that the eigenvectors corresponding to the top largest eigenvalues are typically not very localized. This indicates that latent factors extracted from them have a more pervasive effect on banks in providing credit to the firms.

Interestingly, building Planar Maximally Filtered Graphs (PMFG) from the correlations driven different eigenmodes, we find that the local interaction structure between banks changes in different periods. More specifically, when the dominance of a group of banks in one period gradually disappears, the credit market starts to build-up a different structure in the next period in which another group of banks become the new hubs in the backbone of the cross-correlations.

Several directions for future research can be suggested from our present work. First, as we can see in the analysis of eigenvector elements, it is often observed that in a couple of the eigenvectors corresponding to the top largest eigenvalues, a large group of banks tend to have a higher degree of correlations among themselves. Therefore, the clustering behaviours and the community structure of banks in the correlation matrices in the different lending layers should be studied further (e.g. Almog et al. 2015; MacMahon and Garlaschelli 2015). Second, it would also be interesting to examine the effects of the bank characteristics (e.g. bank locations, bank types, balance sheets' information) on the formation of such clusters and communities. Last but not least, another important direction for future research is to analyse the multilayer architecture of the interdependencies between banks, e.g. to

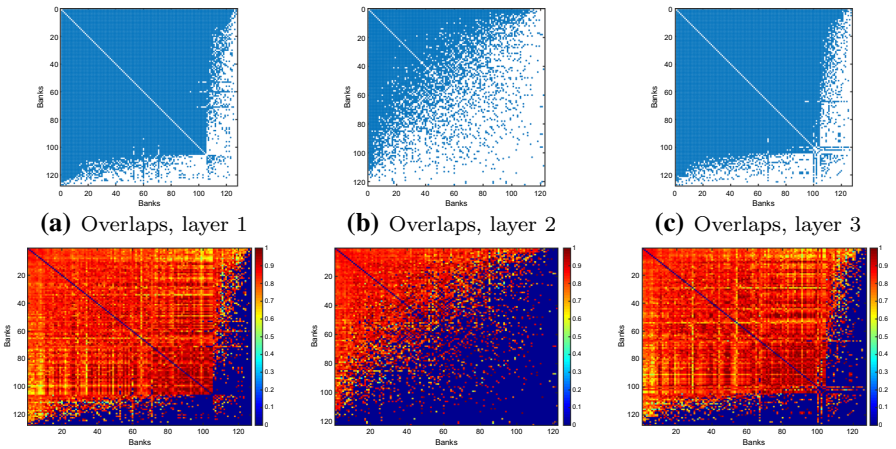
consider together various layers including the network of loan portfolio overlaps, the network of loan portfolio correlations, the correlation network of bank stock returns, and the physical interbank trading network (e.g. Brunetti et al. 2015; Montagna and Lux 2017). We believe that this direction will help to gain a deeper understanding of the complex structure of interactions within the banking system and between the banking system and the other sectors of the economy.

Appendix

Here we provide additional results for portfolio overlaps and similarities, the fractions of eigenvalues in different ranges, the distributions as well as the evolution of the eigenvector elements, and the more detailed correlation network structure between banks based on the method of PMFG. We also report the main results for the binary version of the bank-firm credit network of Japan.

Portfolio Overlaps and Similarities

In this part, we briefly summarize the main results for portfolio overlaps and portfolio similarities between banks. Recall that in each lending layer, each element of the binary version of the bank-bank projection matrix A^{B-B} explained in Eq. (3) implies whether a pair of banks share at least one common borrower. Meanwhile, each element of the similarity matrix S^{B-B}, which is based on the generalized Jaccard index defined in Eq. (1), indicates the degree of structural similarity between two loan portfolios. Figure 10 demonstrates the binary version of the bank-



(a) Overlaps, layer 1 **(b)** Overlaps, layer 2 **(c)** Overlaps, layer 3

(d) Similarity, layer 1 **(e)** Similarity, layer 2 **(f)** Similarity, layer 3

Fig. 10 Binary overlaps and portfolio similarities between banks in the three layers, in 2012. Banks are sorted in descending order of the eigenvector centrality of the adjacency matrix obtained from A^{B-B}. Panels **a** to **c** The binary version of the bank-bank projection matrices in the three layers. Each blue point in these panels indicates a pair of banks have at least one common borrower. Panels **d** to **f** The similarity matrices in the three layers. The colorbar indicates the range [0, 1] for the elements of the similarity matrices .(Colour figure online)

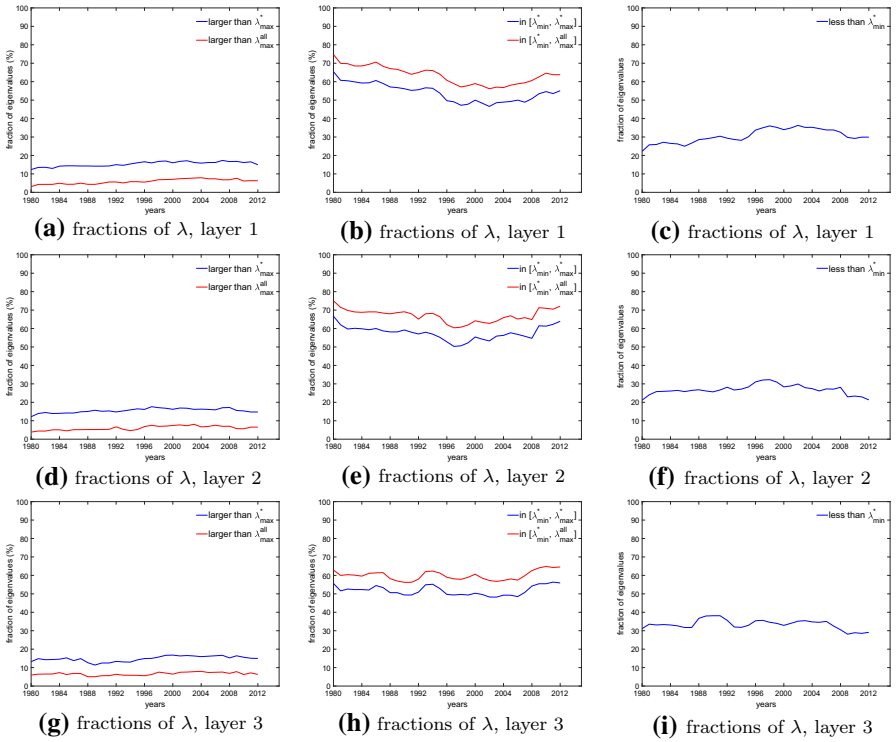


Fig. 11 Fractions of eigenvalues in different ranges for the three layers. Panels **a**, **d**, **g** Fractions of eigenvalues larger than the considered upper bound λ^*_{max} (blue lines) or λ^{all}_{max} (red lines) for layers 1, 2, 3, respectively. Panels **b**, **e**, **h** Fractions of eigenvalues lie within $[\lambda^*_{min}, \lambda^*_{max}]$ (blue lines) or $[\lambda^*_{min}, \lambda^{all}_{max}]$ (red lines) for layers 1, 2, 3, respectively. Panels **c**, **f**, **i** Fractions of eigenvalues less than the lower bound λ^*_{min} for layers 1, 2, 3, respectively .(Colour figure online)

bank projection matrices (panels (a) to (c)) and the similarity matrices (panels (d) to (f)) in the three lending layers, for the data in 2012 as an example. We observe that many pairs of banks have overlaps in their loan portfolios, especially in the layer 1 and layer 3. Furthermore, the loan portfolios also tend to have a relatively high degree of similarities.[15]

Fractions of Eigenvalues in Different Ranges

Figure 11 shows the fractions of eigenvalues in different ranges, i.e. larger than the considered upper bound (λ^*_{max} or $\lambda^{all}_{max} = \max(\lambda^*_{max}, \lambda^1_{max}, \lambda^2_{max})$), within the interval from the lower bound to the considered upper bound ($[\lambda^*_{min}, \lambda^*_{max}]$ or $[\lambda^*_{min}, \lambda^{all}_{max}]$), and less than the lower bound λ^*_{min}. Although the fractions change over the years, in

[15] To illustrate the similarity matrices, for the sake of convenience, we still use the same order of banks as in the analysis of binary overlaps in Fig. 10a–c, i.e. banks are sorted in descending order of the eigenvector centrality (e.g. see Bonacich (2007)) of the adjacency matrix obtained from A^{B-B}. We do not conclude any nature of the relationship between that order and the elements of the portfolio matrix S^{B-B}.

general we find that a majority of the eigenvalues lie within the random bulk or are smaller than the lower bound λ^*_{min}. In contrast, only a small subset of the top eigenvalues persistently exceed the upper bound λ^*_{max} or λ^{all}_{max} over the whole period.

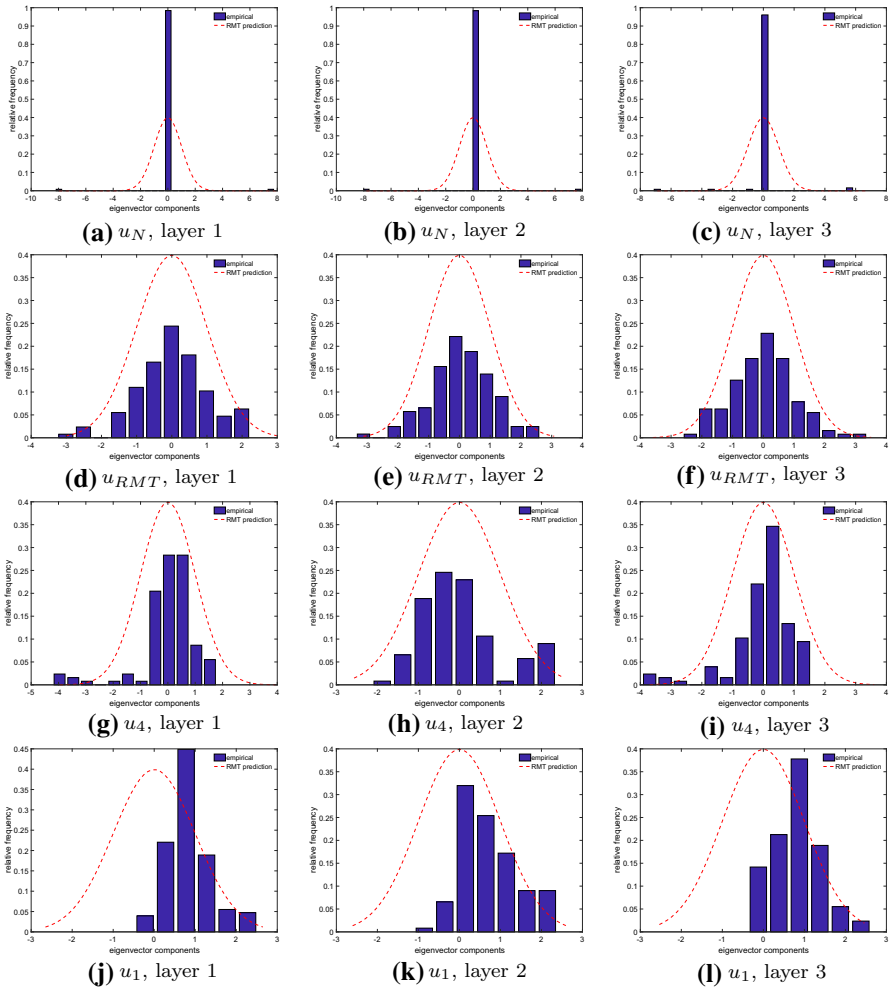
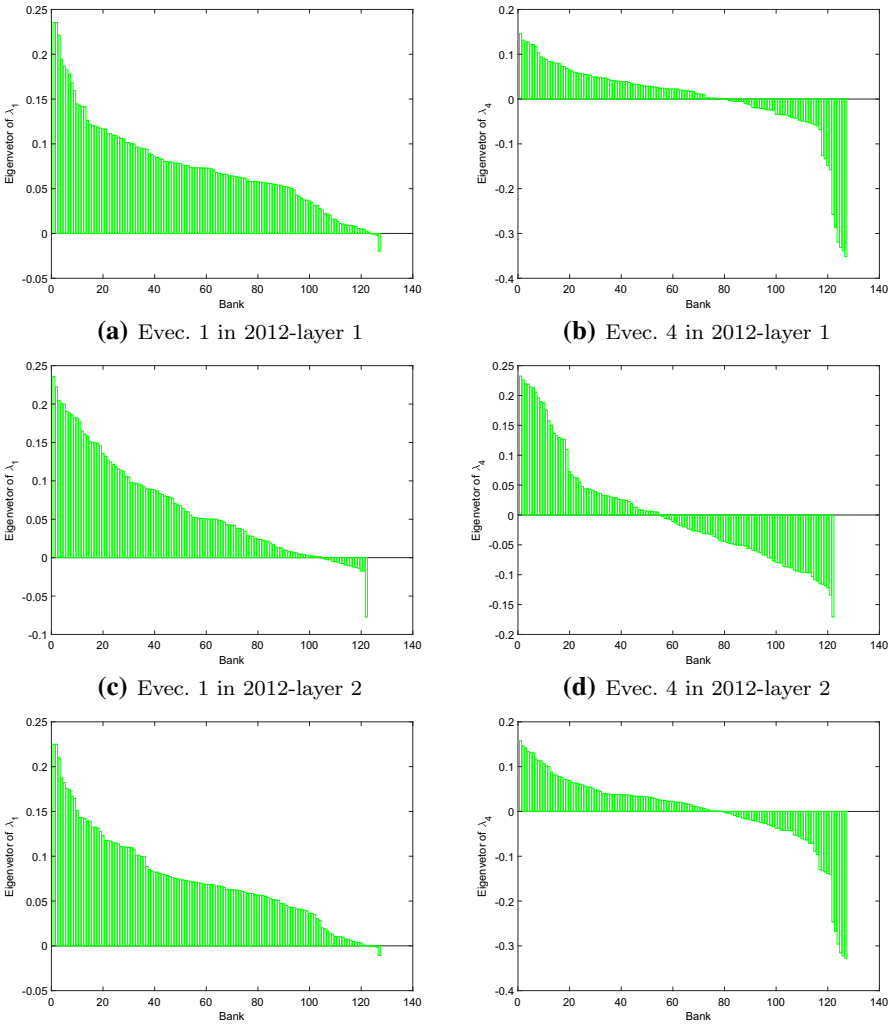


Fig. 12 Distributions of the eigenvector elements of four selected eigenvalues of C, in 2012, in the three layers. Panels **a** to **c** show the distribution of the eigenvector elements of u_N. Panels **d** to **f** show the distribution of the elements of an eigenvector in the random bulk. Panels **g** to **i** show the distribution of the eigenvector elements of u_4. Panels **j** to **l** show the distribution of the eigenvector elements of u_1



(a) Evec. 1 in 2012-layer 1 **(b)** Evec. 4 in 2012-layer 1

(c) Evec. 1 in 2012-layer 2 **(d)** Evec. 4 in 2012-layer 2

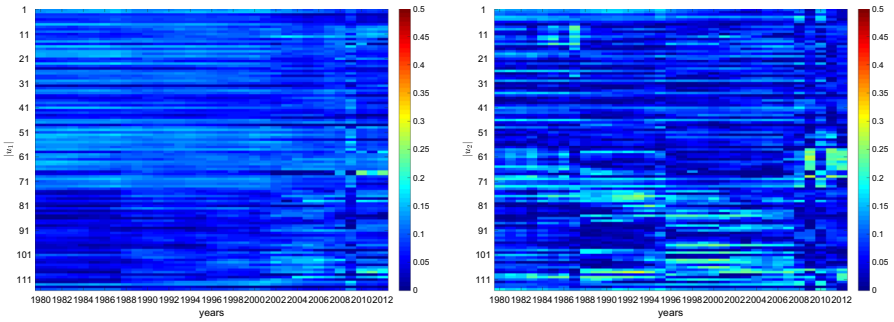
(e) Evec. 1 in 2012-layer 3 **(f)** Evec. 4 in 2012-layer 3

Fig. 13 The eigenvectors (sorted in descending order) of λ_1 and λ_4, in the three layers, in 2012. Without loss of generality, we fix the signs of eigenvector elements such that their sum is non-negative. Panels **a**, **b** show the eigenvectors of λ_1 and λ_4 in the layer 1. Panels **c**, **d** show the eigenvectors of λ_1 and λ_4 in the layer 2. Panels **e**, **f** show the eigenvectors of λ_1 and λ_4 in the layer 3

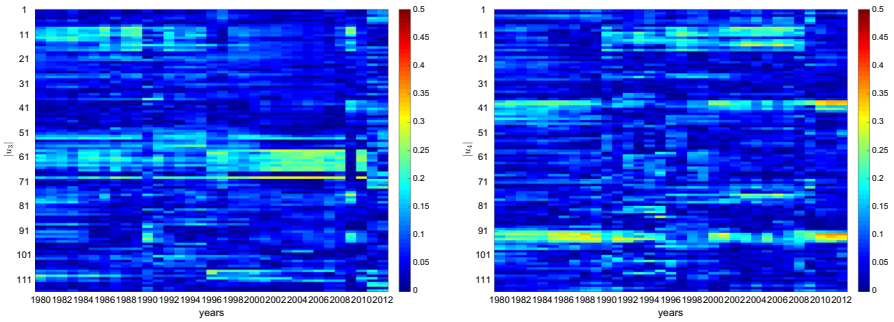
Additional Results for Eigenvector Elements

Distribution of Eigenvector Elements

In Fig. 12, we compare the distributions of the elements of the four selected eigenvectors of C, i.e. the eigenvector u_N associated with the smallest eigenvalue



(a) Elements of u_1 **(b)** Elements of u_2



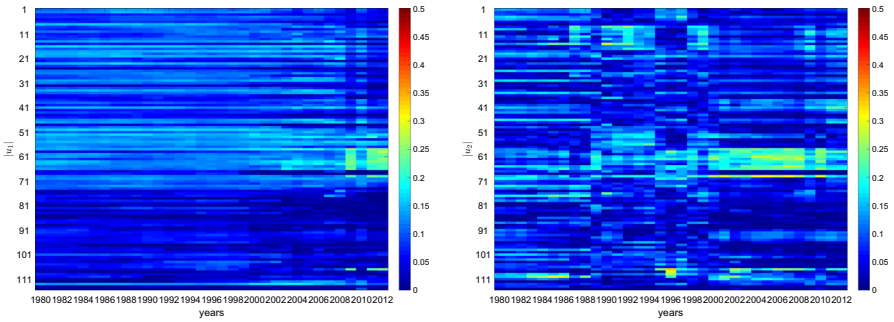
(c) Elements of u_3 **(d)** Elements of u_4

Fig. 14 The evolution of the elements of the eigenvectors corresponding to λ_1 (panel (**a**)), λ_2 (panel (**b**)), λ_3 (panel (**c**)), and λ_4 (panel (**d**)), for 115 banks activating in all years, in the layer 1

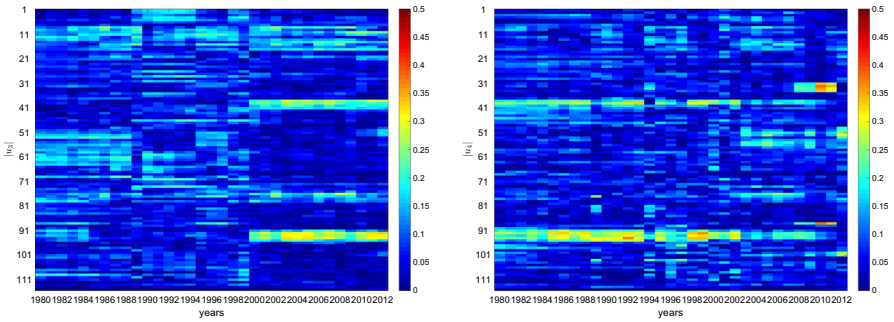
λ_N, the eigenvector u_{RMT} corresponding to a representative eigenvalue located in the random bulk, the eigenvector u_4 associated with λ_4, and the eigenvector u_1 corresponding to the largest eigenvalue λ_1. Note that for the illustration purpose, in this Figure we re-normalize the eigenvector elements such that $\sum_{j=1}^{N} u_i(j)^2 = N$ $(\forall i)$. We can see that u_N has only few significant elements, revealing that this eigenvector is very localized. In addition, the distributions of the elements of u_1 and u_4 are different from the distribution of the elements of u_{RMT}.

Localization Behavior

As two examples, with one for a more localized behavior of eigenvector elements and another for a less localized behavior of eigenvector elements, Fig. 13 shows the entries of u_1 and u_4 in descending order. We can see that in the first and third layers, the homogeneity degree is much stronger in u_1 than in u_4. However, it should be emphasized that this does not implies that all banks contribute equally to the factor extracted from λ_1 and u_1. Indeed, as shown in Fig. 13a, c and e, we still observe a certain level of heterogeneity among banks.



(a) Elements of u_1 **(b)** Elements of u_2

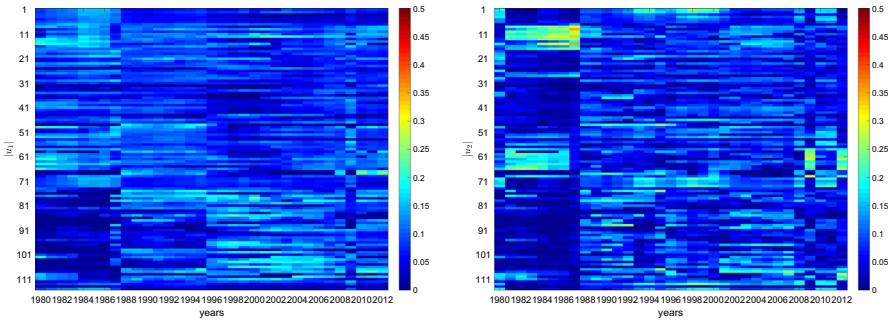


(c) Elements of u_3 **(d)** Elements of u_4

Fig. 15 The evolution of the elements of the eigenvectors corresponding to λ_1 (panel (**a**)), λ_2 (panel (**b**)), λ_3 (panel (**c**)), and λ_4 (panel (**d**)), for 115 banks activating in all years, in the layer 2

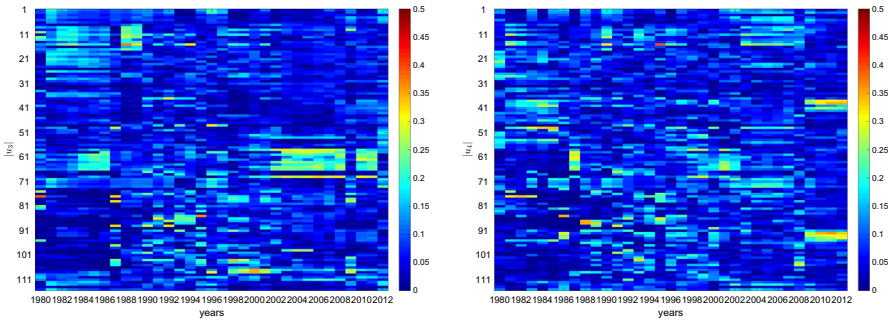
Evolution of Eigenvector Elements

From Figs. 14, 15, 16, we show the evolution of the eigenvectors (in absolute terms) associated with the top four largest eigenvalues in the three lending layers. Here we focus on 115 banks that activate in the bank-firm credit market of Japan in all years from 1980 to 2012. As we can see, some banks correspond to a higher level eigenvector elements in certain sub-periods, but overall, in almost all years, the distribution ranges of these elements are not very wide. This result is consistent with those obtained from the analysis of IPR.



(a) Elements of u_1

(b) Elements of u_2



(c) Elements of u_3

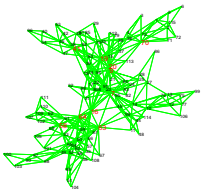
(d) Elements of u_4

Fig. 16 The evolution of the elements of the eigenvectors corresponding to λ_1 (panel (**a**)), λ_2 (panel (**b**)), λ_3 (panel (**c**)), and λ_4 (panel (**d**)), for 115 banks activating in all years, in the layer 3

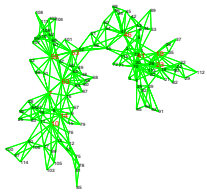
Additional Results for PMFG Graphs Over the Years

In the following we provide additional results for the evolution of PMFG graphs based on the matrix $C^{m,g} = (C^m + C^g)$, which captures the effects of significant factors (i.e. the global common factor as well as the "groups" factors) on the correlations between banks in lending to firms. Again, we focus on 115 banks activating in the bank-firm credit market of Japan in all the years from 1980 to 2012. For the sake of conciseness, here we only show the PMFG graphs in the layer 1 (the total lending layer).[16] In each panel of Figs. 17 and 18, the green links indicate significant pairwise correlations retained in the PMFG graph. The numbers represent banks, and those in red color are the banks that have highest degrees in the PMFG network. We can see that the correlation structure between banks changes in different periods. In particular, when the dominant role of a group of banks in one period gradually vanishes, the credit market starts to build-up a different structure in the next period in which another group of banks will become the new hubs in the backbone of the cross-correlations.

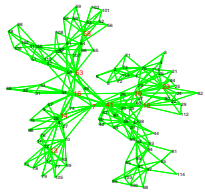
[16] However, the results for the other layers are available from the author upon request.



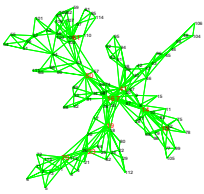
(a) "market" and "groups" modes, in 1980



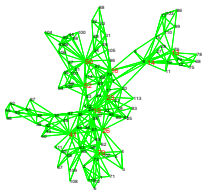
(b) "market" and "groups" modes, in 1981



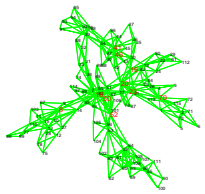
(c) "market" and "groups" modes, in 1982



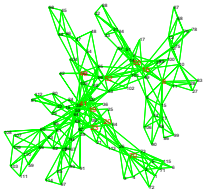
(d) "market" and "groups" modes, in 1983



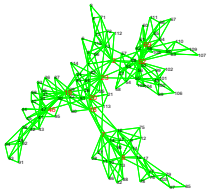
(e) "market" and "groups" modes, in 1984



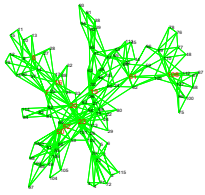
(f) "market" and "groups" modes, in 1985



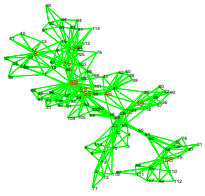
(g) "market" and "groups" modes, in 1986



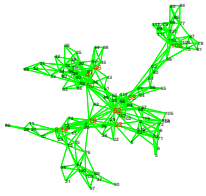
(h) "market" and "groups" modes, in 1987



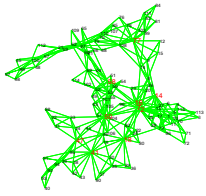
(i) "market" and "groups" modes, in 1988



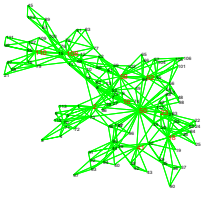
(j) "market" and "groups" modes, in 1989



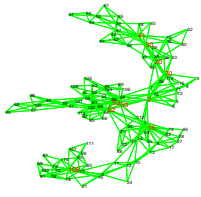
(k) "market" and "groups" modes, in 1990



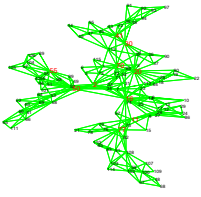
(l) "market" and "groups" modes, in 1991



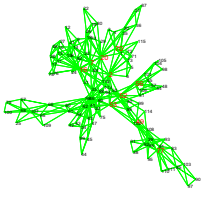
(m) "market" and "groups" modes, in 1992



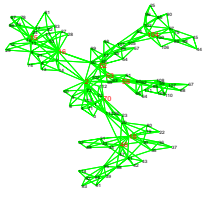
(n) "market" and "groups" modes, in 1993



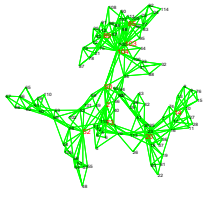
(o) "market" and "groups" modes, in 1994



(p) "market" and "groups" modes, in 1995

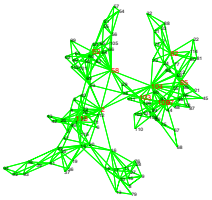


(q) "market" and "groups" modes, in 1996

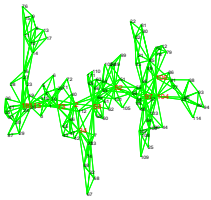


(r) "market" and "groups" modes, in 1997

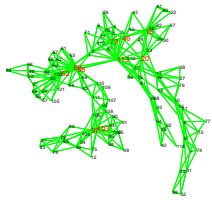
◀**Fig. 17** The interaction structure between banks for the correlation matrix $C^{m,g}$ obtained with the PMFG graph, in layer 1, from 1980 to 1997



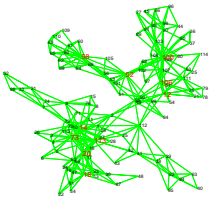
(a) "market" and "groups" modes, in 1998



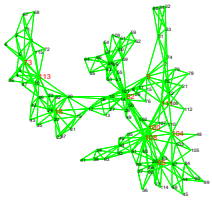
(b) "market" and "groups" modes, in 1999



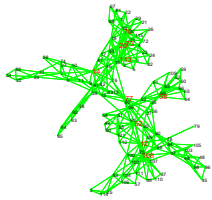
(c) "market" and "groups" modes, in 2000



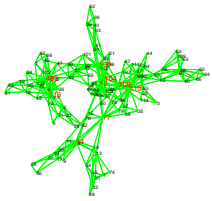
(d) "market" and "groups" modes, in 2001



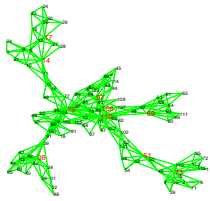
(e) "market" and "groups" modes, in 2002



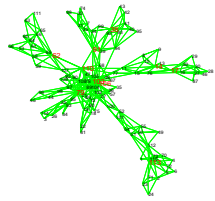
(f) "market" and "groups" modes, in 2003



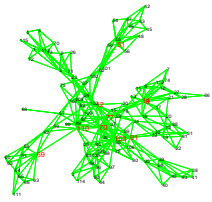
(g) "market" and "groups" modes, in 2004



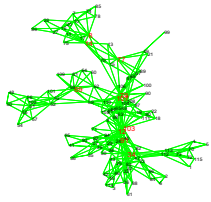
(h) "market" and "groups" modes, in 2005



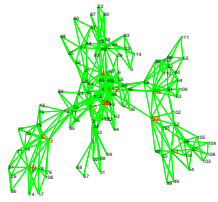
(i) "market" and "groups" modes, in 2006



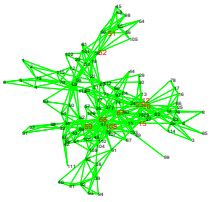
(j) "market" and "groups" modes, in 2007



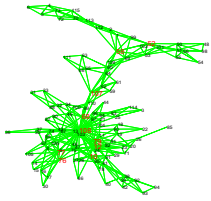
(k) "market" and "groups" modes, in 2008



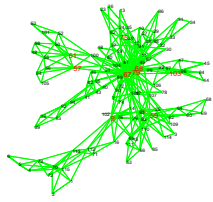
(l) "market" and "groups" modes, in 2009



(m) "market" and "groups" modes, in 2010

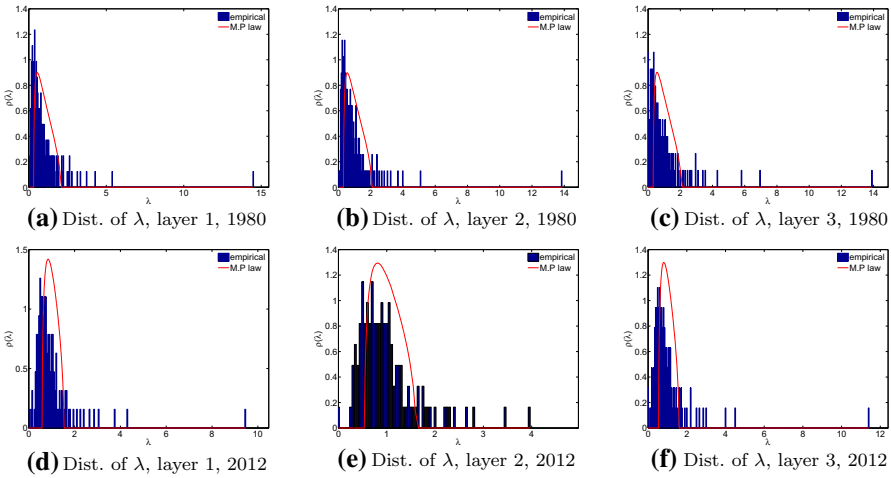


(n) "market" and "groups" modes, in 2011



(o) "market" and "groups" modes, in 2012

Fig. 18 The interaction structure between banks for the correlation matrix $C^{m,g}$ obtained with the PMFG graph, in layer 1, from 1998 to 2012



(a) Dist. of λ, layer 1, 1980 **(b)** Dist. of λ, layer 2, 1980 **(c)** Dist. of λ, layer 3, 1980

(d) Dist. of λ, layer 1, 2012 **(e)** Dist. of λ, layer 2, 2012 **(f)** Dist. of λ, layer 3, 2012

Fig. 19 Binary data: Distribution of eigenvalues of C^{bin}, compared with RMT, in the three layers, in 1980 and 2012. Panels **a**–**c** show the distributions of of λ in 1980, and panels **d**–**f** show the distributions of of λ in 2012. In all panels, the red curve represents the spectral distribution explained by the Marchenko-Pastur law .(Colour figure online)

Binary Analysis

Using the similar methods, we can also analyse the binary version of the bank-firm credit relationships. Interestingly, in general, our results suggest that this version also contains genuine information about the structure and dynamics of the cross-correlation matrix of banks' loan portfolios.

Recalling that under the binary case, for each pair (i, j), we have

$$\begin{cases} a_{i,j}^{bf} = 1 & \text{if bank } i \text{ lends to firm } j, \\ a_{i,j}^{bf} = 0 & \text{otherwise.} \end{cases} \tag{24}$$

We then define the matrix $\mathcal{X}_{bin} = \{\mathrm{X}_{i,j,bin}\}_{N\mathrm{x}N_F}$ with each element is given by

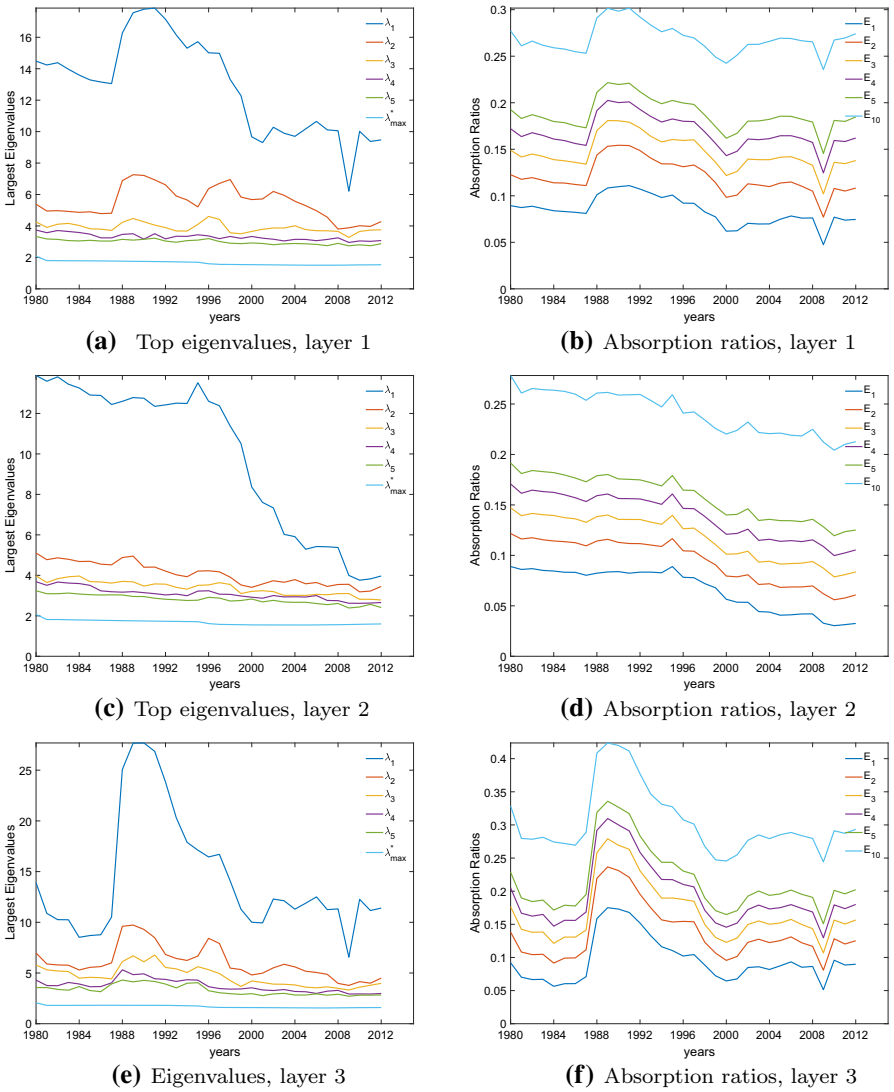
$$\mathrm{X}_{i,j,bin} = \frac{a_{i,j}^{bf} - \langle a_{i,j}^{bf} \rangle}{\sigma(a_{i,j}^{bf})}, \tag{25}$$

where for each bank i, $\langle a_{i,j}^{bf} \rangle$ and $\sigma(a_{i,j}^{bf})$ are the average and the standard deviation of $a_{i,j}^{bf}$ over all N_F firms, respectively. The correlation matrix C^{bin} of the binary data is then defined as

$$C^{bin} = \{C_{ij}^{bin}\}_{N\mathrm{x}N} = \frac{1}{N_F} \mathcal{X}_{bin} \mathcal{X}_{bin}^{\mathcal{T}}, \tag{26}$$

with $\mathcal{X}_{bin}^{\mathcal{T}}$ is the transposition matrix of $\mathcal{X}_{bin}$.

For the sake of conciseness, in the following we will only summarize the main results for the analysis of the binary data. First, we find that a group of largest



(a) Top eigenvalues, layer 1 **(b)** Absorption ratios, layer 1

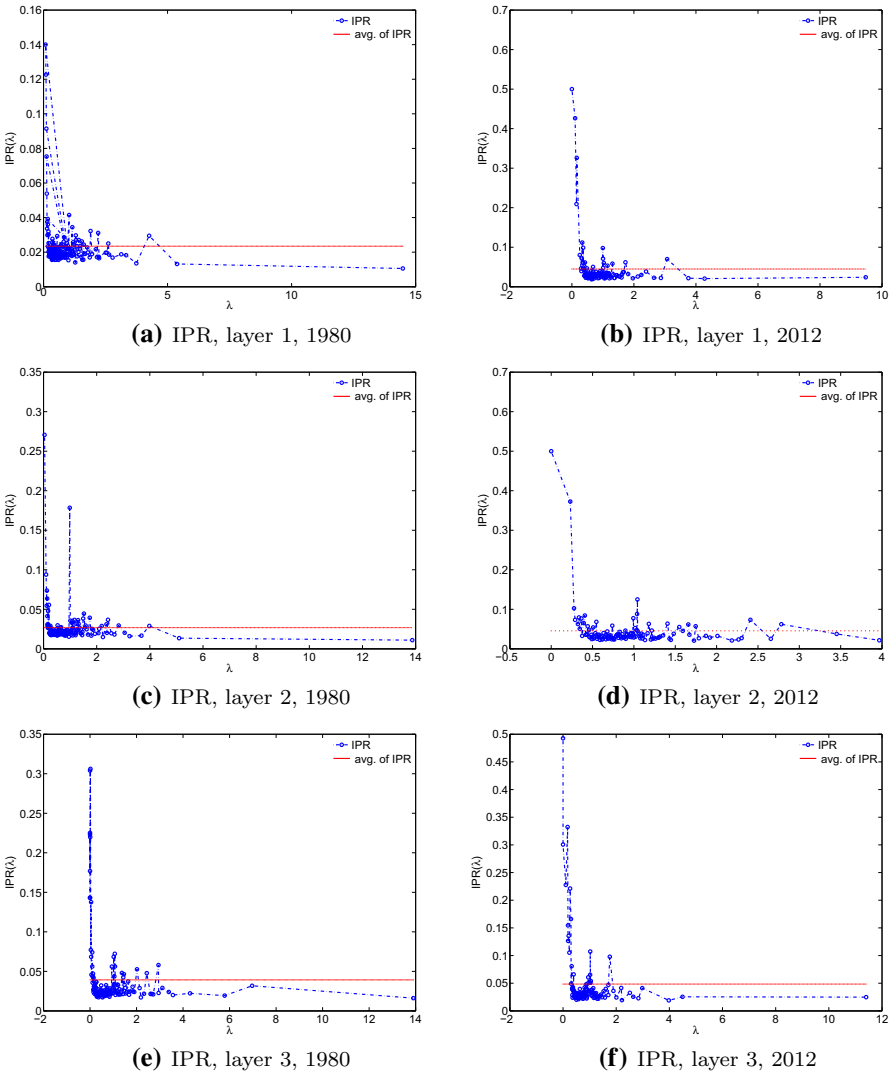
(c) Top eigenvalues, layer 2 **(d)** Absorption ratios, layer 2

(e) Eigenvalues, layer 3 **(f)** Absorption ratios, layer 3

Fig. 20 Binary data: Evolution of the top 5 largest eigenvalues and the associated absorption ratios for C^{bin} in the three layers. The panels on the left show the evolution of the top 5 largest eigenvalues. These eigenvalues are always larger than the upper bound λ^*_{max} implied by RMT. The three panels on the right show their associated absorption ratios as well as the absorption ratio of the top 10 largest eigenvalues

eigenvalues (typically from 10 % to 15 %) always deviate from the random bulk predicted by RMT (see Fig. 19). In addition, we often observe that more than half of the eigenvalues lie between the two RMT bounds λ^*_{min} and λ^*_{max}.

Compared to the weighted data, we also find a similar evolution dynamics for the top largest eigenvalues and their absorption ratios, although in general the factors extracted from these eigenvalues contribute slightly less to the total variance in the



(a) IPR, layer 1, 1980 **(b)** IPR, layer 1, 2012

(c) IPR, layer 2, 1980 **(d)** IPR, layer 2, 2012

(e) IPR, layer 3, 1980 **(f)** IPR, layer 3, 2012

Fig. 21 Binary data: Inverse Participation Ratios of the eigenvectors of C^{bin} in the three layers, in 1980 and 2012. Panels **a**, **b** show IPR versus λ in the layer 1, in 1980 and 2012. Panels **c**, **d** show IPR versus λ in the layer 2, in 1980 and 2012. Panels **e**, **f** show IPR versus λ in the layer 3, in 1980 and 2012. The red lines stand for the averages of $\{\text{IPR}_k\}_{k=1}^{k=N}$

binary data than in the weighted data (e.g. see Fig. 20, in comparison to Fig. 4 in the main text).

Furthermore, as demonstrated in Fig. 21, the localization behavior still emerges in the eigenvectors corresponding to the smallest eigenvalues, as their IPR ratios are often much higher than the average of IPR. Besides this, again many banks contribute to the eigenvectors corresponding to the top largest eigenvalues,

signalling that the "market" mode as well as some of the "groups" modes extracted from the binary version of the credit network also have a wide effect.

Acknowledgements I am indebted to the editors and anonymous reviewers. I am grateful to Prof. Thomas Lux, for valuable comments and discussions that helped to improve the paper. I also thank participants to the different conferences and seminars where earlier versions of this paper were presented. These include: the Conference on Complex System 2018 (CCS 2018) in Thessaloniki, Greece, September 24–28, 2018; the Research Seminars at the Chair of Monetary Economics and International Finance, the University of Kiel, Germany. All remaining errors are my own.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review, 105*(2), 564–608.

Alfaro, L., García-Santana, M., & Moral-Benito, E. (2018). On the direct and indirect real effects of credit supply shocks. CEPR Discussion Papers 12794, CEPR Discussion Papers.

Allen, F., & Gale, D. (2000). Financial contagion. *Journal of Political Economy, 108*(1), 1–33.

Almog, A., Besamusca, F., MacMahon, M., & Garlaschelli, D. (2015). Mesoscopic community structure of financial markets revealed by price and sign fluctuations. *PLOS ONE, 10*(7), 1–16.

Alvarez-Socorro, A. J., Herrera-Almarza, G. C., & González-Díaz, L. A. (2015). Eigencentrality based on dissimilarity measures reveals central nodes in complex networks. *Scientific Reports, 5,* 17095.

Amiti, M., & Weinstein, D. E. (2018). How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data. *Journal of Political Economy, 126*(2), 525–587.

Aoyama, H. (2014). Systemic risk in Japanese credit network. In H. Aoyama, B. K. Chakrabarti, A. Chakraborti, A. Ghosh, & F. Abergel (Eds.), *Econophysics of Agent-Based Models* (pp. 219–228). Berlin: Springer International Publishing.

Aoyama, H., Stefano, B., & Fujiwara, Y. (2013). DebtRank Analysis of the Japanese credit network. RIETI Discussion Paper Series 13-E-087, The Research Institute of Economy, Trade and Industry, Japan.

Aste, T., Shaw, W., & Matteo, T. D. (2010). Correlation structure and dynamics in volatile markets. *New Journal of Physics, 12*(8), 085009.

Azariadis, C. (2018). Credit cycles and business cycles. *Federal Reserve Bank of St Louis Review, 100*(1), 45–71.

Bai, Z., Liu, H., & Wong, W. K. (2009). Enhancement of the applicability of markowitz's portfolio optimization by utilizing random matrix theory. *Mathematical Finance, 19*(4), 639–667.

Bandyopadhyay, J. N., & Jalan, S. (2007). Universality in complex networks: Random matrix analysis. *Physical Review E, 76,* 026109.

Barigozzi, M., Brownlees, C., Gallo, G. M., & Veredas, D. (2014). Disentangling systematic and idiosyncratic dynamics in panels of volatility measures. *Journal of Econometrics, 182*(2), 364–384.

Bassett, W. F., Chosak, M. B., Driscoll, J. C., & Zakrajšek, E. (2014). Changes in bank lending standards and the macroeconomy. *Journal of Monetary Economics, 62,* 23–40.

Beale, N., Rand, D. G., Battey, H., Croxson, K., May, R. M., & Nowak, M. A. (2011). Individual versus systemic risk and the Regulator's Dilemma. *Proceedings of the National Academy of Sciences, 108*(31), 12647–12652.

Bentolila, S., Jansen, M., & Jiménez, G. (2018). When credit dries up: Job losses in the great recession. *Journal of the European Economic Association, 16*(3), 650–695.

Bianconi, G. (2018). *Multilayer networks: Structure and function*. Oxford: Oxford University Press.

Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics, 104*(3), 535–559.

Biroli, G., & Bouchaud, J. P. (2007). Potters M (2007) Extreme value problems in random matrix theory and other disordered systems. *Journal of Statistical Mechanics Theory and Experiment, 07,* P07019.

Bodnar, T., Mazur, S., & Okhrin, Y. (2013). On the exact and approximate distributions of the product of a wishart matrix with a normal vector. *Journal of Multivariate Analysis, 122,* 70–81.

Bodnar, T., Mazur, S., & Okhrin, Y. (2014). Distribution of the product of singular wishart matrix and normal vector. Theory of Probability and Mathematical Statistics pp. 1–14.

Bodnar, T., Mazur, S., & Podgórski, K. (2016). Singular inverse wishart distribution and its application to portfolio theory. *Journal of Multivariate Analysis, 143,* 314–326.

Bodnar, T., Mazur, S., & Podgórski, K. (2017). A test for the global minimum variance portfolio for small sample and singular covariance. *AStA Advances in Statistical Analysis, 101*(3), 253–265.

Bodnar, T., Mazur, S., Muhinyuza, S., & Parolya, N. (2018). On the product of a singular wishart matrix and a singular gaussian vector in high dimension. *Theory of Probability and Mathematical Statistics, 99,* 37–50.

Bodnar, T., Mazur, S., Podgórski, K., & Tyrcha, J. (2019). Tangency portfolio weights for singular covariance matrix in small and large dimensions: Estimation and test theory. *Journal of Statistical Planning and Inference, 201,* 40–57.

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks, 29*(4), 555–564.

Bouchaud, J. P., & Potters, M. (2009). Financial applications of random matrix theory: A short review. ArXiv e-prints.

Bouchaud, J. P., Laloux, L., Miceli, M. A., & Potters, M. (2007). Large dimension forecasting models and random singular value spectra. *European Physical Journal B, 55*(2), 201–207.

Bowman, D., Cai, F., Davies, S., & Kamin, S. (2015). Quantitative easing and bank lending: Evidence from Japan. *Journal of International Money and Finance, 57,* 15–30.

Brummitt, C. D., D'Souza, R. M., & Leicht, E. A. (2012). Suppressing cascades of load in interdependent networks. *Proceedings of the National Academy of Sciences, 109*(12), E680–E689.

Brunetti, C., Harris, J., Mankad, S., & Michailidis, G. (2015). Interconnectedness in the interbank market. Finance and Economics Discussion Series 2015-90, Board of Governors of the Federal Reserve System (U.S.).

Brunnermeier, M. K., & Schnabel, I. (2016). *Bubbles and Central Banks: Historical Perspectives* (pp. 493–562). Cambridge: Cambridge University Press.

Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., & Havlin, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature, 464*(7291), 1025–1028.

Bülbül, D., & Lambert, C. (2012). Credit portfolio modelling and its effect on capital requirements. Discussion Papers 11/2012, Deutsche Bundesbank.

Caccioli, F., Shrestha, M., Moore, C., & Farmer, J. D. (2014). Stability analysis of financial contagion due to overlapping portfolios. *Journal of Banking and Finance, 46,* 233–245.

Caccioli, F., Farmer, J. D., Foti, N., & Rockmore, D. (2015). Overlapping portfolios, contagion, and financial stability. *Journal of Economic Dynamics and Control, 51,* 50–63.

Cai, J., Eidam, F., Saunders, A., & Steffen, S. (2018). Syndication, interconnectedness, and systemic risk. *Journal of Financial Stability, 34,* 105–120.

Chamberlain, G., & Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica, 51*(5), 1281–1304.

Cingano, F., Manaresi, F., & Sette, E. (2016). Does credit crunch investment down? New evidence on the real effects of the bank-lending channel. *The Review of Financial Studies, 29*(10), 2737–2773.

Corsi, F., Marmi, S., & Lillo, F. (2016). When micro prudence increases macro risk: The destabilizing effects of financial innovation, leverage, and diversification. *Operations Research, 64*(5), 1073–1088.

Daly, J., Crane, M., & Ruskin, H. (2008). Random matrix theory filters in portfolio optimisation: A stability and risk assessment. *Physica A Statistical Mechanics and Its Applications, 387*(16), 4248–4260.

de Carvalho, J. X., Jalan, S., & Hussein, M. S. (2009). Deformed gaussian-orthogonal-ensemble description of small-world networks. *Phys Rev E, 79,* 056222.

de Castro Miranda, R. C., & Tabak, B. M. (2013). Contagion risk within firm-bank bivariate networks. Working Papers Series 322, Central Bank of Brazil, Research Department.

De Masi, G., & Gallegati, M. (2012). Bank-firms topology in Italy. *Empirical Economics, 43*(2), 851–866.

De Masi, G., Fujiwara, Y., Gallegati, M., Greenwald, B., & Stiglitz, J. E. (2011). An analysis of the Japanese credit network. *Evolutionary and Institutional Economics Review, 7*(2), 209–232.

Di Matteo, T., Pozzi, F., & Aste, T. (2010). The use of dynamical networks to detect the hierarchical organization of financial market sectors. *The European Physical Journal B, 73*(1), 3–11.

Dumitriu, I., & Johnson, T. (2016). The Marcenko-Pastur law for sparse random bipartite biregular graphs. *Random Structures and Algorithms, 48*(2), 313–340.

Fenech, J. P., Vosgha, H., & Shafik, S. (2015). Loan default correlation using an Archimedean copula approach: A case for recalibration. *Economic Modelling, 47,* 340–354.

Forni, M., & Lippi, M. (2001). The generalized dynamic factor model: Representation theory. *Econometric Theory, 17*(6), 1113–1141.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *The Review of Economics and Statistics, 82*(4), 540–554.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2015). The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association, 100*(471), 830–840.

Forni, M., Hallin, M., Lippi, M., & Zaffaroni, P. (2017). Dynamic factor models with infinite-dimensional factor space: Asymptotic analysis. *Journal of Econometrics, 199*(1), 74–92.

Fricke, D. (2012). Trading strategies in the overnight money market: Correlations and clustering on the e-MID trading platform. *Physica A Statistical Mechanics and Its Applications, 391*(24), 6528–6542.

Fricke, D. (2016). Has the banking system become more homogeneous? Evidence from banks' loan portfolios. *Economics Letters, 142,* 45–48.

Fricke, D., & Roukny, T. (2020). Generalists and specialists in the credit market. *Journal of Banking and Finance, 112,* 105335.

Fujiwara, Y., Aoyama, H., Ikeda, Y., Iyetomi, H., & Souma, W. (2009). Structure and temporal change of the credit network between banks and large firms in Japan. Economics: The Open-Access. *Open-Assessment E-Journal, 3,* 1–18.

Gambetti, L., & Musso, A. (2017). Loan supply shocks and the business cycle. *Journal of Applied Econometrics, 32*(4), 764–782.

Gertler, M., & Kiyotaki, N. (2010). Chapter 11 - financial intermediation and credit policy in business cycle analysis. Handbook of Monetary Economics, vol 3, Elsevier, pp 547–599.

Gewers, F. L., Ferreira, G. R., de Arruda, H. F., Silva, F. N., Comin, C. H., Amancio, D. R., & da F Costa, L. (2018). Principal component analysis: A natural approach to data exploration. arXiv:CoRR abs/1804.02502.

Greenwood, R., Landier, A., & Thesmar, D. (2015). Vulnerable banks. *Journal of Financial Economics, 115*(3), 471–485.

Guhr, T., Müller-Groeling, A., & Weidenmüller, H. A. (1998). Random-matrix theories in quantum physics: Common concepts. *Physics Reports, 299*(4), 189–425.

Haldane, A. G., & May, R. M. (2011). Systemic risk in banking ecosystems. *Nature, 469,* 351–355.

Himino, R. (2021). *The Japanese Banking Crisis*. Singapore: Palgrave Macmillan.

Hoshi, T., & Kashyap, A. (2000). The Japanese banking crisis: Where did it come from and how will it end? In: NBER Macroeconomics Annual 1999, Volume 14, National Bureau of Economic Research, Inc, pp 129–212.

Huang, X., Gao, J., Buldyrev, S. V., Havlin, S., & Stanley, H. E. (2011). Robustness of interdependent networks under targeted attack. *Physical Review E, 83,* 065101.

Huang, X., Vodenska, I., Havlin, S., & Stanley, H. E. (2013). Cascading failures in bi-partite graphs: Model for systemic risk propagation. *Scientific Reports, 3,* 1219.

Iyetomi, H., Nakayama, Y., Yoshikawa, H., Aoyama, H., Fujiwara, Y., Ikeda, Y., & Souma, W. (2011). What causes business cycles? Analysis of the Japanese industrial production data. *Journal of the Japanese and International Economies, 25*(3), 246–272.

Jalan, S. (2009). Spectral analysis of deformed random networks. *Physical Review E, 80,* 046101.

Jiang, X., Chen, T., & Zheng, B. (2014). Structure of local interactions in complex financial dynamics. *Scientific Reports, 4,* 5321.
Jiang, X. F., & Zheng, B. (2012). Anti-correlation and subsector structure in financial systems. *EPL (Europhysics Letters), 97*(4), 48006.
Jolliffe, I. (1986). *Principal Component Analysis*. Berlin: Springer Verlag.
Jorda, O., Schularick, M. H. P., & Taylor, A. M. (2013). When credit bites back. *Journal of Money, Credit and Banking, 45*(s2), 3–28.
Kawai, M., & Takagi, S. (2011). Why was Japan hit so hard by the global financial crisis? In: The impact of the economic crisis on East Asia, Edward Elgar Publishing, chap 7.
Khwaja, A. I., & Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review, 98*(4), 1413–42.
Kim, D. H., & Jeong, H. (2005). Systematic analysis of group identification in stock markets. *Physical Review E, 72,* 046133.
Kiyotaki, N., & Moore, J. (1997). Credit cycles. *Journal of Political Economy, 105*(2), 211–248.
Kritzman, M., Li, Y., Page, S., & Rigobon, R. (2011). Principal components as a measure of systemic risk. *The Journal of Portfolio Management, 37*(4), 112–126.
Kyriakopoulos, F., Thurner, S., Puhr, C., & Schmitz, S. W. (2009). Network and eigenvalue analysis of financial transaction networks. *The European Physical Journal B, 71*(4), 523.
Laloux, L., Cizeau, P., Bouchaud, J. P., & Potters, M. (1999). Noise dressing of financial correlation matrices. *Physical Review Letters, 83,* 1467–1470.
Laloux, L., Cizeau, P., Potters, M., & Bouchaud, J. P. (2000). Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance, 03*(03), 391–397.
Lehmann, S., Schwartz, M., & Hansen, L. K. (2008). Biclique communities. *Physical Review E, 78,* 016108.
Levy-Carciente, S., Kenett, D. Y., Avakian, A., Stanley, H. E., & Havlin, S. (2015). Dynamical macroprudential stress testing using network theory. *Journal of Banking and Finance, 59,* 164–181.
Lillo, F., & Pirino, D. (2015). The impact of systemic and illiquidity risk on financing with risky collateral. *Journal of Economic Dynamics and Control, 50,* 180–202.
Liu, R. R., Li, M., & Jia, C. X. (2016). Cascading failures in coupled networks: The critical role of node-coupling strength across networks. *Scientific Reports, 6,* 35352.
Luu, D. T., & Lux, T. (2018). Identifying patterns in the bank-sector credit network of Spain. *Journal Of Network Theory in Finance, 4*(1), 1–38.
Luu, D. T., & Lux, T. (2019). Multilayer overlaps and correlations in the bank-firm credit network of Spain. *Quantitative Finance, 19*(12), 1953–1974.
Lux, T. (2016). A model of the topology of the bank - firm credit network and its role as channel of contagion. *Journal of Economic Dynamics and Control, 66,* 36–53.
Lux, T. (2020). On the distribution of links in financial networks: Structural heterogeneity and functional form. *Empirical Economics, 58*(3), 1019–1053.
Lux, T., Luu, D. T., & Yanovski, B. (2020). An analysis of systematic risk in worldwide econonomic sentiment indices. *Empirica, 47,* 909–928.
MacMahon, M., & Garlaschelli, D. (2015). Community detection for correlation matrices. *Physical Review X, 5,* 021006.
Mantegna, R. N. (1999). Hierarchical structure in financial markets. *European Physical Journal B, 11*(1), 193–197.
Marotta, L., Miccichè, S., Fujiwara, Y., Iyetomi, H., Aoyama, H., Gallegati, M., & Mantegna, R. N. (2015). Bank-firm credit network in Japan: An analysis of a bipartite network. *PLOS ONE, 10*(5), 1–18.
Marotta, L., Miccichè, S., Fujiwara, Y., Iyetomi, H., Aoyama, H., Gallegati, M., & Mantegna, R. N. (2016). Backbone of credit relationships in the Japanese credit market. *EPJ Data Science, 5*(1), 10.
Marčenko, V. A., & Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik, 1*(4), 457.
Massara, G. P., di Matteo, T., & Aste, T. (2015). Network filtering for Big Data: Triangulated Maximally Filtered Graph. arXiv:CoRR abs/1505.02445.
Matteo, B., & Marc, H. (2016a). Generalized dynamic factor models and volatilities: Recovering the market volatility shocks. *The Econometrics Journal, 19*(1), C33–C60.
Matteo, B., & Marc, H. (2016b). A network analysis of the volatility of high dimensional financial series. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 66*(3), 581–605.

May, R. M., & Arinaminpathy, N. (2010). Systemic risk: The dynamics of model banking systems. *Journal of The Royal Society Interface, 7*(46), 823–838.

Meng, H., Xie, W. J., Jiang, Z. Q., Podobnik, B., Zhou, W. X., & Stanley, H. (2014). Systemic risk and spatiotemporal dynamics of the US housing market. *Scientific Reports, 4,* 3655.

Montagna, M., & Lux, T. (2017). Contagion risk in the interbank market: A probabilistic approach to cope with incomplete structural information. *Quantitative Finance, 17*(1), 101–120.

Nier, E., Yang, J., Yorulmazer, T., & Alentorn, A. (2007). Network models and financial stability. *Journal of Economic Dynamics and Control, 31*(6), 2033–2060 (**Tenth Workshop on Economic Heterogeneous Interacting Agents.**).

Ormerod, P. (2008). Random matrix theory and macro-economic time-series: An illustration using the evolution of business cycle synchronisation, 1886–2006. *Economics the Open-Access Open-Assessment E-Journal, 2,* 1–10.

Ormerod, P., & Mounfield, C. (2000). Random matrix theory and the failure of macro-economic forecasts. *Physica A: Statistical Mechanics and its Applications, 280*(3), 497–504.

Pappas, D., & Kaimakamis, G. (2010). Optimal portfolio selection with singular covariance matrix. *International Mathematical Forum, 5,* 2305–2318.

Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L. A., & Stanley, H. E. (1999). Universal and nonuniversal properties of cross correlations in financial time series. *Physical Review Letters, 83,* 1471–1474.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L., & Stanley, H. (2000). A random matrix theory approach to financial cross-correlations. *Physica A Statistical Mechanics and Its Applications, 287*(3), 374–382.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T., & Stanley, H. E. (2002). Random matrix approach to cross correlations in financial data. *Physical Review E, 65,* 066126.

Pool, V. K., Stoffman, N., & Yonker, S. E. (2015). The people in your neighborhood: Social interactions and mutual fund portfolios. *Journal of Finance, 70*(6), 2679–2732.

Posen, A. S. (2003). It takes more than a bubble to become Japan. Working Paper Series WP03-9, Peterson Institute for International Economics.

Potestio, R., Caccioli, F., & Vivo, P. (2009). Random matrix approach to collective behavior and bulk universality in protein dynamics. *Physical Review Letters, 103,* 268101.

Pozzi, F., Aste, T., Rotundo, G., & Matteo, T. D. (2008). Dynamical correlations in financial systems. In: Proceedings of SPIE, vol 6802, pp. 6802–6813.

Pukthuanthong, K., & Roll, R. (2009). Global market integration: An alternative measure and its application. *Journal of Financial Economics, 94*(2), 214–232.

Quintana, D., García-Rodríguez, S., Cincotti, S., & Isasi, P. (2015). Combining RMT-based filtering with time-stamped resampling for robust portfolio optimization. *International Journal of Computational Intelligence Systems, 8*(5), 874–885.

Reis, S. D. S., Hu, Y., Babino, A., Andrade, J. S., Jr., Canals, S., Sigman, M., & Makse, H. A. (2014). Avoiding catastrophic failure in correlated networks of networks. *Nature Physics, 10,* 762–767.

Sharifi, S., Crane, M., Shamaie, A., & Ruskin, H. (2004). Random matrix theory for portfolio optimization: A stability approach. *Physica A Statistical Mechanics and Its Applications, 335*(3), 629–643.

Shen, J., & Zheng, B. (2009). Cross-correlation in financial dynamics. *EPL (Europhysics Letters), 86*(4), 48005.

Smart, A. G., Amaral, L. A. N., & Ottino, J. M. (2008). Cascading failure and robustness in metabolic networks. *Proceedings of the National Academy of Sciences, 105*(36), 13223–13228.

Srivastava, M. (2003). Singular wishart and multivariate beta distributions. The Annals of Statistics 31.

Tao, T., & Vu, V. (2012). Random covariance matrices: Universality of local statistics of eigenvalues. *The Annals of Probability, 40*(3), 1285–1315.

Tran, L. V., Vu, V. H., & Wang, K. (2013). Sparse random graphs: Eigenvalues and eigenvectors. *Random Structures and Algorithms, 42*(1), 110–134.

Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences, 102*(30), 10421–10426.

Uechi, L., Akutsu, T., Stanley, H. E., Marcus, A. J., & Kenett, D. Y. (2015). Sector dominance ratio analysis of financial markets. *Physica A Statistical Mechanics and Its Applications, 421,* 488–509.

Wang, D., Podobnik, B., Horvatić, D., & Stanley, H. E. (2011). Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices. *Physical Review E, 83,* 046121.

Wang, G. J., Xie, C., & Stanley, H. E. (2018). Correlation structure and evolution of world stock markets: Evidence from pearson and partial correlation-based networks. *Computational Economics, 51*(3), 607–635.

Yoshikawa, H., Aoyama, H., Fujiwara, Y., & Iyetomi, H. (2015). Deflation/inflation dynamics: Analysis based on micro prices . CIRJE F-Series CIRJE-F-959, CIRJE, Faculty of Economics, University of Tokyo.

Zheng, Z., Podobnik, B., Feng, L., & Li, B. (2012). Changes in cross-correlations as an indicator for systemic risk. *Scientific Reports, 2,* 888.