ORIGINAL PAPER

A dynamic network model to measure exposure concentration in the Austrian interbank market

Juraj Hledik[1] · Riccardo Rastelli[2]

Abstract

Motivated by an original financial network dataset, we develop a statistical methodology to study non-negatively weighted temporal networks. We focus on the characterization of how nodes (i.e. financial institutions) concentrate or diversify the weights of their connections (i.e. exposures) among neighbors. The approach takes into account temporal trends and nodes' random effects. We consider a family of nested models on which we define and validate a model-selection procedure that can identify those models that are relevant for the analysis. We apply the methodology to an original dataset describing the mutual claims and exposures of Austrian financial institutions between 2008 and 2011. This period allows us to study the results in the context of the financial crisis in 2008 as well as the European sovereign debt crisis in 2011. Our results highlight that the network is very heterogeneous with regard to how nodes send, and in particular receive edges. Also, our results show that this heterogeneity does not follow a significant temporal trend, and so it remains approximately stable over the time span considered.

✉ Riccardo Rastelli
riccardo.rastelli@ucd.ie

1   European Commission's Joint Research Centre, Ispra, Italy

2   Present Address: School of Mathematics and Statistics, University College Dublin, Dublin, Ireland

1 Motivation

Recently, the EU was hit by two major financial crises. In 2008, the problems started initially in the US subprime mortgage market and were partially caused by lax regulation and overly confident debt ratings. The source of the European sovereign debt crisis in 2011, however, was most likely private debt arising from property bubble and resulting in government bailouts. The lack of a common fiscal union in the EU did not help with the situation, which resulted in the European central bank providing cheap loans to maintain a steady cash flow between EU banks. During these turbulent times, European banks were facing high levels of uncertainty. It was not clear which counter party would remain solvent in the foreseeable future and even sovereign bonds were no longer considered a safe option. In the face of these unfavorable conditions, the banks were forced to reconsider their interbank investments and readjust their portfolios in order to account for the change in the economic situation.

With the goal of contributing to the discussion on diversification and concentration of interbank exposures, we focus our attention on an original dataset of the Austrian interbank market between the spring of 2008 and autumn of 2011. Namely, we introduce a dynamic network model to quantify exposure concentration of individual banks and of the market overall. We accomplish this by extending the current literature on latent variable network models to a framework for compositional data that evolve over time. This framework can potentially provide bank-dependent measures of systemic risk, as well as a global measure of the overall level of systemic risk in the market. We resort to an intuitive modeling of a single network homogeneity (drift) parameter which we use to capture the homogeneity over time, as well as node random effects parameters that capture *exposure concentration* and *attractiveness* of banks, individually. Our model is specifically designed for instances where a network needs to be characterized by a single evolving variable, or when one is interested in obtaining a model-based quantitative measurement of the inter-temporal development of network homogeneity. While our approach is strongly motivated by our original financial application, it also provides a new methodology extension that can be used to analyze any temporal and compositional network data, where the main interest lies in studying the concentration/diversification of edge weights. This contributes to the literature on the statistical analysis of weighted networks, and in particular to the literature on networks indicating proportions, which is a rather unexplored topic. The code associated to our methodology is available from the public repository (Rastelli and Hledik 2021).

1.1 Motivating example

In a financial context, a change in a financial network structure can have far-reaching and non-trivial consequences. To illustrate this fact further, consider a hypothetical financial network of four institutions (banks) represented by nodes and their mutual financial exposures (debt) represented by edges. In this simple example, connections
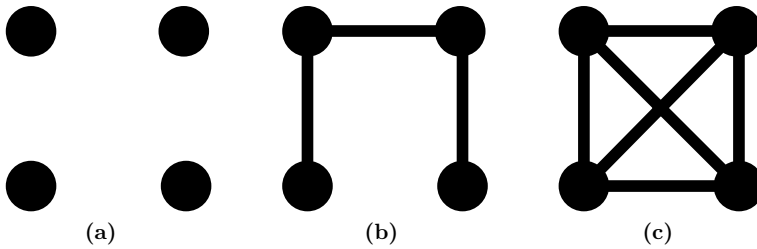
Fig. 1 Different loan network structures on a set of four banks

are symmetric and every bank splits its investment among its neighbors equally. Furthermore, banks are required by a regulator to always keep a capital buffer to account for unexpected withdrawals, unfavorable economic conditions and other factors. Therefore, we assume that an institution remains safe unless it loses at least half of its investments in other institutions. If that happens, the institution gets bankrupt and it might further negatively affect other banks in the network. To see how network structure affects the overall stability, consider a case where one of these four banks gets affected by an exogenous shock such that it has to declare bankruptcy. In such case, its neighbors will not get their respective investment and might suffer the same fate, putting their own neighbors in danger. This contagious behavior is dependent on how the banks are linked together, which illustrates the importance of structure when addressing questions on systemic importance and financial stability.

For the hypothetical case of four banks, there are up to 11 topologically different network structures that can possibly occur: a subset of these are shown in Fig. 1. In the case shown in Fig. 1a, there is no danger of contagion since there are no edges to propagate shocks. An analogical result follows from the network shown in Fig. 1c, where a failure of one node is not sufficient to take down the rest because every other institution only loses one third of its investment. Problems arise in intermediately connected systems such as Fig. 1b, where an initial shock may wipe out the whole system.

This basic example hints at a much more complex issue of network stability that has been extensively studied by financial regulators in the past two decades. It highlights that the level of exposure concentration in a system plays a crucial role in determining its stability and that assessment of this trait for observed networks can prove challenging. In this paper, we address this impasse, introducing a statistical model specifically designed to measure the concentration of exposures for every bank in the system, hence obtaining a measure for one of the facets of systemic risk.

2 Related literature

This paper is connected to two important strands of research. On the one hand, we contribute to the literature on statistical analysis of networks, since we extend the current available methodologies, based on latent variables, that can be used to analyze non-negatively weighted temporal networks. On the other hand, we contribute to the established literature on systemic risk and financial networks. This area of

research has often focused on the stability of financial systems as well as the possibility of contagious bankruptcies similar to our simple example above. Research papers on this subject have been published by both academics in finance as well as market regulators.[1] Our proposed research borrows from and contributes to both fields, ultimately creating a novel perspective on financial systemic risk.

Within the statistics literature, our paper contributes to the research on latent variable network modeling. Prominent contributions include the foundational paper on latent space models by Hoff et al. (2002), later extended to the dynamic framework by Sarkar and Moore (2006), and the latent stochastic blockmodels (Nowicki and Snijders 2001) extended to a dynamic framework by Yang et al. (2011), Xu and Hero (2014) and Matias and Miele (2017), among others. These latent variable models possess a number of desirable theoretical features, as illustrated in Rastelli et al. (2016) and Daudin et al. (2008) for the latent space model and stochastic blockmodel, respectively.

Our approach also shares a number of similarities with other recent papers that apply a latent variable framework on various types of dynamic network data. These include, among others, Friel et al. (2016), where the authors introduce a dynamic latent space model to measure the financial stability of the Irish Stock Exchange; Sewell and Chen (2016), who introduce a latent space model for dynamic weighted networks; and Matias et al. (2018), where the authors propose a dynamic extension of the stochastic blockmodel. Our model could be seen as a variant of the works by Westveld and Hoff (2011) and Sewell (2018), which deal with temporal latent variable weighted networks, to the context of compositional data. Differently from all these works using latent variable frameworks, our approach relies on a new model which is designed to measure and study the systemic risk associated to a financial system. We point out that our approach also has similarities to Dirichlet regression models for compositional data, as per Minka (2000), van der Merwe (2019) and references therein; however, differently from these works, we apply compositional data analysis to the context of network analysis.

Within the finance literature, one of the earliest papers on the topic of systemic risk was the work of Allen and Gale (2000), who have shown that the structure of the interbank market is important for the evaluation of possible contagious bankruptcies. Later on, Gai and Kapadia (2010) extended their work from a simple model of four institutions to a financial network of an arbitrary size. Other notable papers on systemic risk include, for example, Glasserman and Young (2016) or Acemoglu et al. (2015), while Upper (2011) provides an excellent survey of regulatory-published scientific reports on the subject. With respect to the questions on diversification, we refer the reader to Elliott et al. (2014) and Frey and Hledik (2018), where a nontrivial relationship between diversification, exposure concentration and contagious defaults is presented, or to Goncharenko et al. (2018), where

---

[1] This includes various national central banks as well as the European Central Bank and the FED. Additional research has been undertaken by the Bank for International Settlements or the International Monetary Fund.

banks endogenously choose their level of diversification given the network structure in an equilibrium setting.

For a better understanding of the topic of systemic risk and financial networks and of its apparent interdisciplinary nature, see the literature review of Caccioli et al. (2018). For a better picture of how these tools can potentially be used for a macro-prudential risk assessment, see Gai and Kapadia (2019) or Battiston and Martinez-Jaramillo (2018). The role of diversification and exposure concentration and its intricate effect on financial stability is addressed in Bardoscia et al. (2017). This paper also underlines the importance of network topology and its effects on overall stability of the financial system.

Our paper shares similar goals with these works, and we further add to these papers by introducing a new generative mechanism and a modeling framework where exposure concentration and homogeneity of the system can be studied inter-temporally.

Lastly, we also contribute to the literature on the stability of the Austrian inter-bank market. Related works in this area include Elsinger et al. (2006), Puhr et al. (2012) and Boss et al. (2004) who have looked at possible contagious effects and descriptive statistics of the Austrian financial network. Compared to these contributions, one novelty of our work is that we are able to consider the temporal dependency of the network, thus providing an appreciation of the changes in the structure of the network over time.

3 Data and exploratory analysis

We use an original dataset obtained by the Austrian National Bank which contains quarterly observations of the Austrian interbank market for a period of four years (from spring of 2008 until autumn of 2011). More precisely, the dataset contains aggregated mutual claims between any two of $N = 800$ Austrian banks for all relevant quarters (2008Q1–2011Q4), resulting in 16 observations of the financial network. All of the banks considered exist throughout the whole period.

In order to comply with the privacy rules of the Austrian National Bank, the data is anonymized such that the true identities of banks in the system are hidden and replaced by non-descriptive IDs. Moreover, we are unable to observe the absolute values of banks' mutual claims, only their scaled equivalents (relative to the highest exposure value, independently for each time frame). As a consequence, as per privacy protection, the magnitude of claims is effectively not comparable across time. This motivated us to seek an approach that could bypass this limitation: for the purposes of our model, the true values of the claims are not required, since our framework only uses their relative size. In addition, we illustrate in Appendix A a procedure that allows us to approximate the true values of the claims up to a proportionality constant: we do not use these estimated quantities in our model, but we use them to reconstruct and approximate a value representing the importance of each institution. In order to better clarify these concepts, we now give a sequence of definitions to set our notation.

3.1 Definitions

A *dynamic network of interbank exposures* is a sequence of graphs where, for each time frame, the nodes correspond to banks and the edges correspond to the connections between them. In particular, the edges are directed and carry positive values indicating the claim on one bank from another. We note that an observed network of interbank exposures between N banks over T time frames may be represented as a collection of T interaction matrices of the same size $N \times N$, as in the following definition:

Definition 3.1 A sequence of true exposures $\mathcal{E} = \{\mathbf{E}^{(t)}\}_{t \in \mathcal{T}}$ defined on the set of nodes $\mathcal{V}$ over the timespan $\mathcal{T}$ consists of a collection of interaction matrices $\mathbf{E}^{(t)} \ \forall t \in \mathcal{T}$ with elements $e_{ij}^{(t)}$ for $t \in \mathcal{T}, i \in \mathcal{V}, j \in \mathcal{V}$, where $e_{ij}^{(t)}$ corresponds to the financial exposure of bank i towards bank j in period t.

We focus on the case where $\mathcal{V} = \{1, \dots, N\}$ and $\mathcal{T} = \{1, \dots, T\}$. In the Austrian interbank market context, the adjacency matrix $\mathbf{E}^{(t)}$ contains the true values of all mutual claims between any two of $N = 800$ Austrian banks at the corresponding time frame. However, as explained earlier, we are unable to observe the true exposures $\mathcal{E}$ due to privacy policy of the Austrian National Bank. By contrast, we adopt a procedure (see Appendix A for full details) that allows us to estimate the following quantities:

Definition 3.2 A *sequence of absolute exposures* $\mathcal{X} = \{\mathbf{X}^{(t)}\}_{t \in \mathcal{T}}$ on the set of nodes $\mathcal{V}$ over the timespan $\mathcal{T}$ has elements defined as follows:

$$x_{ij}^{(t)} \overset{def}{=} \frac{e_{ij}^{(t)}}{\max_{\{k,l\}} e_{kl}^{(1)}} \quad \forall i, j \in \mathcal{V}, \forall t \in \mathcal{T} \tag{1}$$

In other words, the sequence of absolute exposures is simply a scaled version of the non-observable sequence of the true exposures, where every exposure is divided by the value of the first period's largest exposure. Lastly, we define the sequence of relative exposures that our statistical model uses as observed data:

Definition 3.3 A *sequence of relative exposures* $\mathcal{Y} = \{\mathbf{Y}^{(t)}\}_{t \in \mathcal{T}}$ on the set of nodes $\mathcal{V}$ over the time span $\mathcal{T}$ has elements defined as follows:

$$y_{ij}^{(t)} \overset{def}{=} \frac{x_{ij}^{(t)}}{\sum_{k=1}^{N} x_{ik}^{(t)}} \quad \forall i, j \in \mathcal{V}, \forall t \in \mathcal{T} \tag{2}$$

This transformation constricts the edge weights in our networks to a [0, 1] interval, making it easier to work with from the network homogeneity viewpoint. In this network, every nodes' outgoing edge values always sum up to 1. Differently from $\mathcal{X}$,

the data $\mathcal{Y}$ is not estimated and available to us in exact form, regardless of the transformations and anonymizations.

To summarize, $\mathcal{E}$ corresponds to the *non-observed* real value of interbank connections, $\mathcal{X}$ to their estimated scaled version and $\mathcal{Y}$ to the relative interbank connections. Our model only uses $\mathcal{Y}$ as observed data, which in fact corresponds to the only quantities that are available to us in an exact form. We use $\mathcal{X}$ as a proxy to derive a measure of importance of banks. In order to do so, we introduce the bank's *relevance*:

Definition 3.4 The relevance of bank i in time period t is defined as:

$$r_i^{(t)} = \sum_{k \in \mathcal{V}} x_{ik}^{(t)} + \sum_{k \in \mathcal{V}} x_{ki}^{(t)}. \tag{3}$$

Thus, the relevance simply corresponds to the bank's overall sum of its interbank assets and liabilities. Again, the relevance variable is not used in the modeling, instead it is only used to contextualize the results.

With a clear measure of systemic importance, we can now select a subsample of banks with the highest *aggregated relevance* $r_i = \sum_{t=1}^{T} r_i^{(t)}$. This allows us to focus on the interactions of systemically important banks and observe emergence of unique patterns. We use the aggregated relevance measure to create a smaller dataset consisting of the 200 systemically most relevant institutions and their mutual connections. From now on, we shall refer to the full dataset and the reduced dataset as `OeNB 800` and `OeNB 200`, respectively.[2]

3.2 Exploratory data analysis

In order to have a better overview of the data, we have conducted an exploratory analysis. We plot the evolution of the average bank relevance in Fig. 2. As expected from its definition, we see a sharp drop in the second half of 2008 as a direct effect of the financial crisis.

Table 1 and Fig. 3 contain brief descriptive statistics, where one can see the number as well as magnitude of connections as a function of time. The number of connections (2nd column) shows the number of edges in the network as of time t, while the *relative* size (3rd column) is proportional to the overall total exposure, calculated as the ratio between the average absolute exposure in each year, and the highest absolute exposure in the first time observation. We would like to highlight the second and third quarter of 2008, where a drop in the overall total exposure in the economy can be observed. This period corresponds to the financial crisis associated with the failure of Lehman Brothers in the US and the problems stemming from the housing market. Interestingly, in the Austrian interbank market, the overall number

---

[2] Validity of the `OeNB 200` subset can be justified further by examining the overall exposure of top 200 institutions. It turns out that the 200 most systemically relevant banks account for at least 95% of all approximate edge weights in any given time frame (according to the definition of absolute exposures from 3.2).
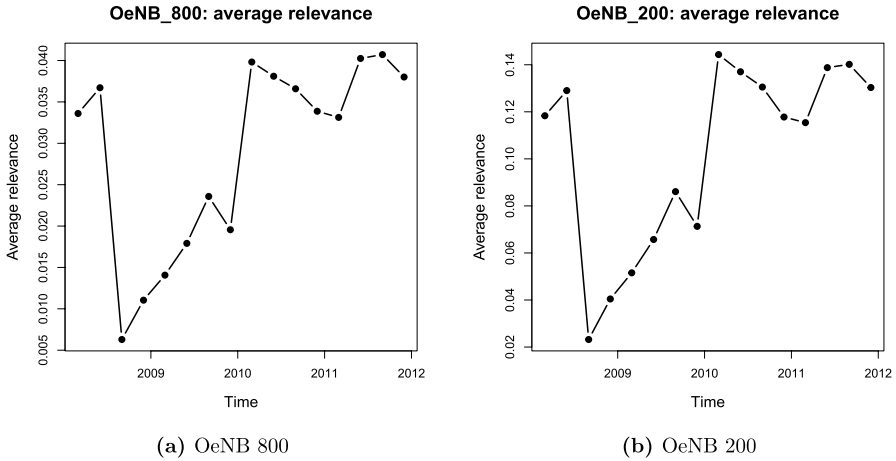
OeNB_800: average relevance

OeNB_200: average relevance

(**a**) OeNB 800

(**b**) OeNB 200

Fig. 2 Bank relevance for the full sample (**a**) and the sample containing only the 200 most relevant banks (**b**), recalculated after the exclusion of the 600 banks
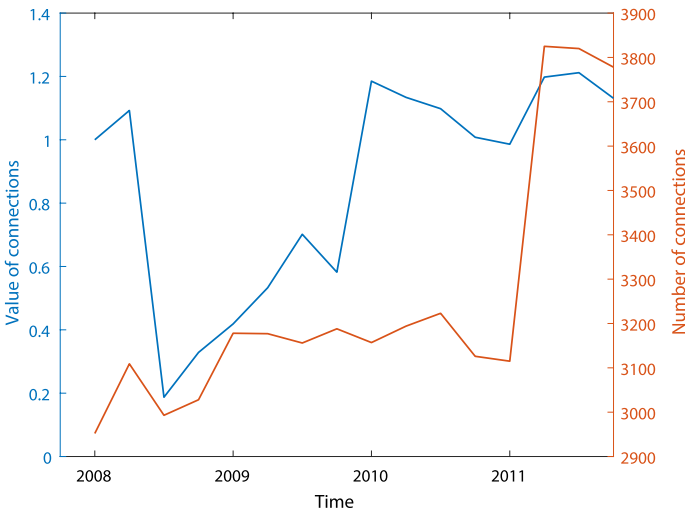

Fig. 3 Number of connections and their relative size in time, for `OeNB 800`

of connections does not seem to be affected by these events as much as their size. This shows that, albeit Austrian banks have reduced their mutual exposures significantly, they were rarely completely cut off. Another important period is during the second and third quarter of 2011, which is roughly when the European sovereign debt crisis started. At a first glance, there does not seem to be much in relation to this event in our data.

Interbank markets are commonly disassortative, i.e. nodes with a low number of neighbors are mostly connected to nodes with high number of neighbors and vice

Table 1 Number of connections and their relative size in time, for `OeNB 800`

Period	No. of connections	Relative size of connections
2008Q1	2952	1.0000
2008Q2	3109	1.0925
2008Q3	2993	0.1873
2008Q4	3028	0.3287
2009Q1	3178	0.4186
2009Q2	3177	0.5329
2009Q3	3156	0.7016
2009Q4	3188	0.5820
2010Q1	3157	1.1851
2010Q2	3194	1.1340
2010Q3	3223	1.0981
2010Q4	3126	1.0080
2011Q1	3115	0.9860
2011Q2	3825	1.1979
2011Q3	3820	1.2118
2011Q4	3778	1.1310

versa, see Hurd (2016). This property in financial networks is quite common and is generally referred to as "core-periphery" structure. Social networks tend to be fundamentally different, since a high number of "hubs" in the network does not necessarily imply a low number of triangles, see for instance (Li et al. 2014; Watts and Strogatz 1998). Financial systems also tend to be very sparse. These same patterns are confirmed in the Austrian interbank market, as shown in Fig. 4.

3.3 Entropy as a measure of concentration

We have observed several interesting patterns in the data which suggest that using a more complex model could produce new insights regarding the evolution of bank exposure concentration. Since the main interest of our research lies in the connectivity patterns of agents in an interbank market, we have also looked at the evolution of an entropy index for this system. For this purpose, we use a standard definition of entropy as follows:

Definition 3.5 The *entropy* $S_i^{(t)}$ of node $i \in \mathcal{V}$ at time $t \in \mathcal{T}$ is defined as:

$$S_i^{(t)} \overset{def}{=} - \sum_{k=1}^{N} y_{ik}^{(t)} \log y_{ik}^{(t)} \tag{4}$$

with the convention that $y \log y = 0$ when $y = 0$.

(a) Adjacency matrix   (b) Plot of a network snapshot
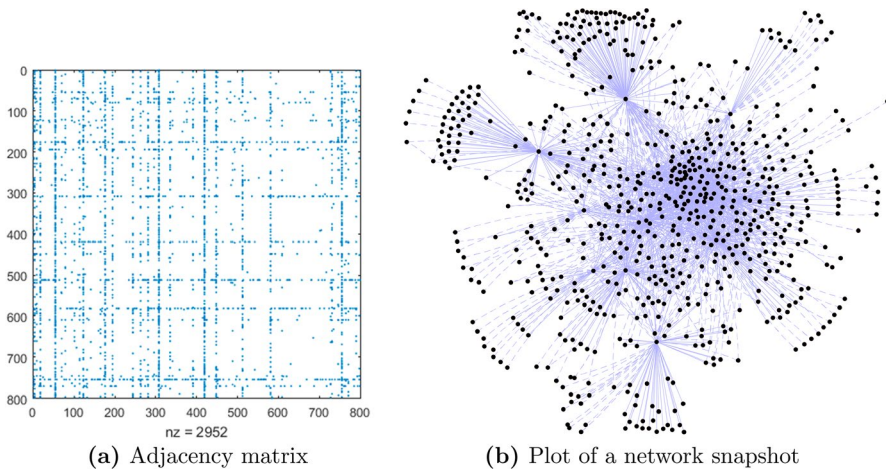
Fig. 4 Adjacency matrix for the first time period, for `OeNB 800`, consisting of 2952 edges represented as dots (**a**) and a graphical representation of the network snapshot for the nodes with at least one connection (**b**)

Speaking more plainly, this quantity describes how an institution distributes its assets among counterparties. A bank with a single debtor would have zero entropy, since its relative exposure is trivially one for that one debtor and zero for all the other banks. With an increased number of debtors with equal exposures, a node's entropy increases and, for a fixed number of debtors, the entropy of a node is maximized when its assets are distributed evenly among neighbors. Ergo, if two nodes have the same number of outgoing connections, one may view the one with a higher entropy as better diversified. We will refer to "better diversified" nodes as ones with lower *exposure concentration* or, alternatively, as the ones with higher *exposure spread*.

In Fig. 5, we plot the change in nodes' entropies in consecutive periods $(S_i^{(t+1)} - S_i^{(t)})$. One can observe an increase in both mean and variance during the second and third quarter of 2011, which corresponds to the sovereign crisis in Europe. At that point, future bailouts of several EU countries were uncertain which might have added to the volatility in the market. Interestingly, no similar effect can be seen during the 2008 crisis. We point out that there are other ways of assessing the temporal evolution of node exposure homogeneity. We have chosen the entropy index for our exploratory analysis, as it constitutes a simple, clean and easily tractable approach, but one could easily turn to other measures, e.g. the Herfindahl index as is common practice in economics literature.

Fig. 5 Distribution of entropy change in time, for `OeNB 800`
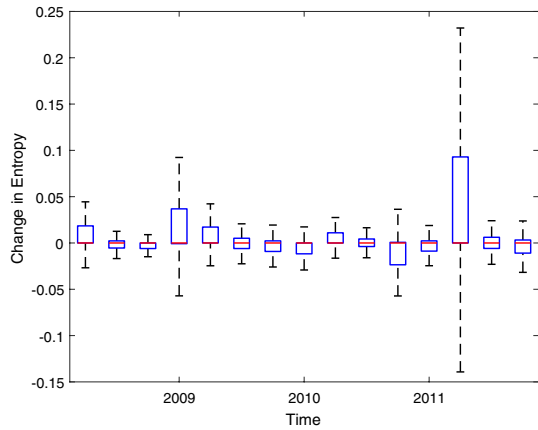
4 The model

4.1 The full model

The observed data are the relative interbank exposures $y_{ij}^{(t)}$ from Definition 3.3. We assume that there are no self-connections, i.e., when not stated otherwise, we always work with $t \in \mathcal{T}$, $i, j \in \mathcal{V}$ and $i \neq j$. Since these exposures are relative, it follows from definition that they satisfy (for all i and t):

$$y_{ij}^{(t)} \in [0, 1] \text{ and } \sum_{j \in \mathcal{V}: j \neq i} y_{ij}^{(t)} = 1. \tag{5}$$

We propose to model the vector $\mathbf{y}_{i \cdot}^{(t)} = \left(y_{i1}^{(t)}, \ldots, y_{iN}^{(t)} \right)$ as a Dirichlet random vector characterized by the parameters $\boldsymbol{\alpha}_{i \cdot}^{(t)} = \left(\alpha_{i1}^{(t)}, \ldots, \alpha_{iN}^{(t)} \right)$, where $\alpha_{ij}^{(t)} > 0$. The values $y_{ii}^{(t)}$ are excluded from the analysis for all $i \in \mathcal{V}$ and $t \in \mathcal{T}$. Following the established standard in latent variable models, the data are assumed to be conditionally independent given the latent parameters $\boldsymbol{\alpha} = \left\{ \alpha_{ij}^{(t)} \right\}_{i,j,t}$. Hence, the model likelihood reads as follows:

$$\mathcal{L}_y(\boldsymbol{\alpha}) = \prod_{t=1}^{T} \prod_{i=1}^{N} \left\{ \frac{\Gamma\left(\sum_j \alpha_{ij}^{(t)} \right)}{\prod_j \Gamma\left(\alpha_{ij}^{(t)} \right)} \prod_j \left[y_{ij}^{(t)} \right]^{\alpha_{ij}^{(t)} - 1} \right\} \tag{6}$$

where, again, j varies in $\mathcal{V}$ and is different from i, and $\Gamma(\cdot)$ denotes the gamma function.

As concerns the $\boldsymbol{\alpha}$ parameters, we separate a trend component from the sender and receiver random effects through the following deterministic representation:

$$\log \left(\alpha_{ij}^{(t)} \right) = \mu^{(t)} + \theta_i + \gamma_j \tag{7}$$

With this formulation, the model parameters $\boldsymbol{\mu} = \left\{\mu^{(t)}\right\}_{t\in\mathcal{T}}$, $\boldsymbol{\theta} = \left\{\theta_i\right\}_{i\in\mathcal{V}}$ and $\boldsymbol{\gamma} = \left\{\gamma_j\right\}_{j\in\mathcal{V}}$ possess a straightforward interpretation, which we illustrate in the next section. Note that the relation between the model parameters and $\boldsymbol{\alpha}$ is deterministic, meaning that $\boldsymbol{\alpha}$ is used only to clarify the notation, and it may be omitted from the model specification.

4.2 Interpretation of model parameters

Before we move to parameter interpretation, we would like to highlight how a symmetric parameter vector $\boldsymbol{\alpha} = \{\alpha, \dots, \alpha\}$ can affect the realizations of the random vector $\mathbf{y} \sim Dir(\boldsymbol{\alpha})$. If the value of α increases, then the variance of the components of the random vector $\mathbf{y}$ tend to decrease. Since the values generated from a Dirichlet distribution lie in an N-dimensional simplex, low variance translates to $y_i \approx 1/N, \forall i \in \mathcal{V}$, e.g. the values are more or less equally distributed (as per their marginal distributions). High variance, however, is obtained when the value of α is small, and it implies that one of the components turns out to be close to one while all the others are close to zero. These two examples closely mimic the high-entropy homogeneous regime and the low-entropy heterogeneous regime introduced in Sect. 3, respectively. Similarly, when the Dirichlet parameter is a non-symmetric random vector $\mathbf{y} \sim Dir(\alpha_1, \dots, \alpha_N)$, an increase in a single parameter component α_j determines a higher expected value in y_j, at the expense of the other elements in $\mathbf{y}$.

In our formulation, the log-additive structure in Eq. (7) decomposes $\boldsymbol{\alpha}$ in three parts. The contribution given by $\mu^{(t)} + \theta_i$ affects all of the components of $\boldsymbol{\alpha}_{i\cdot}^{(t)}$ in a symmetric fashion. Hence, in accordance with our explanation above, these parameters capture the level of homogeneity in the network through a homogeneity trend parameter $\mu^{(t)}$ and a node specific homogeneity random effect θ_i. In financial terms, an increase in either $\mu^{(t)}$ or θ_i corresponds to lower concentration of exposures for bank i at time t, resulting in a more homogeneous network structure. Vice versa, a decrease in $\mu^{(t)}$ or θ_i is linked with an increase in exposure concentration which in turn results in a more heterogeneous network structure.

The interpretation of γ_j is similar. In our context, an increase in γ_j tends to increase the weight of all edges that j receives from its counter parties. Equivalently, one can say that in such case the bank j becomes more *attractive*, in the spirit of other banks concentrating their exposures more towards j.

To summarize, there is a clear way to interpret the main parameters of our model. The parameter $\mu^{(t)}$ indicates the *global homogeneity level* at time frame $t \in \mathcal{T}$, the parameter θ_i characterizes the individual bank i homogeneity level as a sender random effect (its exposure *diversification*), and the parameter γ_j represents the bank j's *attractiveness* as a receiver random effect.

4.3 Nested models

We consider simpler versions of our full model, where some of the elements that compose the $\boldsymbol{\alpha}$s are removed. The models that we consider are defined by different versions of Eq. (7), as follows:

`full:`	$\log\left(\alpha_{ij}^{(t)}\right) = \mu^{(t)} + \theta_i + \gamma_j$
`sender-receiver:`	$\log\left(\alpha_{ij}^{(t)}\right) = c + \theta_i + \gamma_j$
`trend-receiver:`	$\log\left(\alpha_{ij}^{(t)}\right) = \mu^{(t)} + \gamma_j$
`trend-sender:`	$\log\left(\alpha_{ij}^{(t)}\right) = \mu^{(t)} + \theta_i$
`trend:`	$\log\left(\alpha_{ij}^{(t)}\right) = \mu^{(t)}$
`sender:`	$\log\left(\alpha_{ij}^{(t)}\right) = c + \theta_i$
`receiver:`	$\log\left(\alpha_{ij}^{(t)}\right) = c + \gamma_j$
`null:`	$\log\left(\alpha_{ij}^{(t)}\right) = c$

Here, the value c is a parameter that replaces μ when μ is not included in the model, thus absorbing any effect that is homogeneous across all nodes and time frames. The models above represent simpler frameworks whereby each of the elements of our model (exposure diversification trend μ, exposure diversification random effect θ, attractiveness γ) can be removed.

4.4 Bayesian hierarchical structure

We complete our models by introducing the following Bayesian hierarchical structure on the parameters that we have mentioned earlier.

Whenever the drift parameters $\boldsymbol{\mu}$ are present in the model, we assume a random walk process prior on them:

$$\mu_1 \sim \mathcal{N}(0, 1/\tau_\mu), \mu_t = \mu_{t-1} + \eta_t, \quad \forall t > 1,$$

where $\eta_t \sim \mathcal{N}(0, 1/\tau_\eta)$ and $\tau_\eta \sim Gamma(a_\eta, b_\eta)$. The model parameter c follows the same distribution of μ_1. The hyperparameter τ_μ is user-defined and set to a small value to support a wide range of initial conditions. The hyperparameters a_η and b_η are also user-defined and set to small values (0.01) to allow a flexible prior structure.

Whenever the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ are present in the model, they are assumed to be i.i.d. Gaussian variables with:

$$\theta_i \sim \mathcal{N}\left(0, 1/\tau_\theta\right), \tau_\theta \sim \text{Gamma}(a_\theta, b_\theta),$$
$$\gamma_j \sim \mathcal{N}\left(0, 1/\tau_\gamma\right), \tau_\gamma \sim \text{Gamma}(a_\gamma, b_\gamma)$$
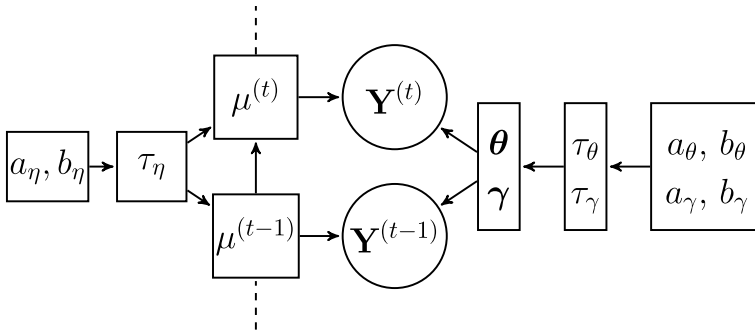
Fig. 6 Graphical representation of model dependencies

Similarly to the other hyperparameters, a_θ, b_θ, a_γ and b_γ are also set to small values (0.01). The arrangement of parameters in Fig. 6 summarizes the dependencies in the `full` model graphically.

5 Parameter estimation

In this section, we focus on the estimation procedure for the `full` model only, noting that the estimation of the nested models follows analogous procedures, whereby the removed parameters are set to zero and are not updated.

Our proposed model has T drift parameters (denoted by $\boldsymbol{\mu}$), N exposure diversification parameters (denoted by $\boldsymbol{\theta}$), N attractiveness parameters (denoted by $\boldsymbol{\gamma}$), and three precision parameters (denoted by $\boldsymbol{\tau}$). We describe in this section a procedure to jointly estimate all of these model parameters.

5.1 Identifiability

The additive structure in Eq. (7) yields a non-identifiable likelihood model. For example, one could define $\tilde{\theta}_i = \theta_i + d$ and $\tilde{\gamma}_j = \gamma_j - d$ for some $d \in \mathbb{R}$ and the likelihood value would be the same for the two configurations, i.e. $\mathcal{L}_y(\boldsymbol{\mu}, \tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\gamma}}) = \mathcal{L}_y(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\gamma})$. One way to deal with such identifiability problem would be to include a penalization through the priors on $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$. One could specify more informative Gaussian priors centered in zero, which would in turn shrink the parameters to be distributed around zero.

However, such approach may also interfere with the results, since the model would not be able to capture the presence of outliers. Hence, we opt for a more commonly used method, and impose the $\boldsymbol{\gamma}$s to sum to zero as expressed through the following constraint:

$$\gamma_1 = -\sum_{j=2}^{N} \gamma_j. \tag{8}$$

This new model, characterized by $T + 2N - 1$ parameters, is now identifiable with respect to its posterior distribution.

5.2 Markov chain Monte Carlo

The posterior distribution associated to our model factorizes as follows:

$$\pi(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\tau}|\mathcal{Y}) \propto \mathcal{L}_{\mathcal{Y}}(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\gamma})\pi(\boldsymbol{\mu}|\tau_\eta)\pi(\boldsymbol{\theta}|\tau_\theta)\pi(\boldsymbol{\gamma}|\tau_\gamma)\times$$
$$\times \pi(\tau_\eta|a_\eta, b_\eta)\pi(\tau_\theta|a_\theta, b_\theta)\pi(\tau_\gamma|a_\gamma, b_\gamma) \tag{9}$$

We adopt a fully Bayesian approach, relying on a Markov chain Monte Carlo algorithm to obtain a random sample from the posterior distribution (9). We use a Metropolis-within-Gibbs sampler that alternates the following steps[3]:

1. Sample μ_s for all $s \in \mathcal{T}$ from the following full-conditional using Metropolis-Hastings with a Gaussian proposal:

$$\pi(\mu_s|\dots) \propto \left\{ \prod_i \Gamma\left(e^{\mu_s}e^{\theta_i}\sum_j e^{\gamma_j} \right) \right\} \left\{ \prod_{i,j} \frac{\left[y_{ij}^{(s)} \right]^{\alpha_{ij}^{(s)}-1}}{\Gamma\left(\alpha_{ij}^{(s)} \right)} \right.$$
$$\left. \cdot \left\{ \exp\left\{ -\frac{\tau_\mu[\mu_s]^2}{2} \right\} \right\}^{\mathbb{1}_{\{s=1\}}} \right\}$$
$$\cdot \left\{ \exp\left\{ -\frac{\tau_\eta[\mu_s - \mu_{s-1}]^2}{2} \right\} \right\}^{\mathbb{1}_{\{s>1\}}}$$
$$\cdot \left\{ \exp\left\{ -\frac{\tau_\eta[\mu_{s+1} - \mu_s]^2}{2} \right\} \right\}^{\mathbb{1}_{\{s<T\}}}. \tag{10}$$

2. Sample θ_k for all $k \in \mathcal{V}$ from the following full-conditional using Metropolis-Hastings with a Gaussian proposal:

$$\pi(\theta_k|\dots) \propto \left\{ \prod_t \Gamma\left(e^{\mu_t}e^{\theta_k}\sum_j e^{\gamma_j} \right) \right\} \left\{ \prod_{t,j} \frac{\left[y_{kj}^{(t)} \right]^{\alpha_{kj}^{(t)}-1}}{\Gamma\left(\alpha_{kj}^{(t)} \right)} \right\} \exp\left\{ -\frac{\tau_\theta}{2}\theta_k^2 \right\}. \tag{11}$$

3. Sample γ_ℓ for all $\ell \in \mathcal{V} \setminus \{1\}$ from the following full-conditional using Metropolis-Hastings with a Gaussian proposal:

---

[3] Note that, in the equations for the parameter updates, the products are defined over the spaces $\mathcal{T}$ and $\mathcal{V}$, with the only restriction that j and ℓ are always different from i. Also, $\mathbb{1}_{\mathcal{A}}$ is equal to 1 if the event $\mathcal{A}$ is true or zero otherwise.

$$\pi(\gamma_\ell|\dots) \propto \left\{ \prod_{t,i} \Gamma\left(e^{\mu_t} e^{\theta_i} \sum_j e^{\gamma_j} \right) \right\} \left\{ \prod_{t,i} \frac{\left[y_{i\ell}^{(t)} \right]^{\alpha_{i\ell}^{(t)} - 1}}{\Gamma\left(\alpha_{i\ell}^{(t)} \right)} \right\}$$

$$\cdot \left\{ \prod_{t,i} \frac{\left[y_{i1}^{(t)} \right]^{\alpha_{i1}^{(t)} - 1}}{\Gamma\left(\alpha_{i1}^{(t)} \right)} \right\} \exp\left\{ -\frac{\tau_\gamma}{2} \gamma_\ell^2 \right\}. \qquad (12)$$

4. Sample τ_η from the following conjugate full-conditional:

$$\pi(\tau_\eta|\dots) \sim Gamma\left(a_\eta + \frac{T-1}{2}, b_\eta + \sum_{t>1} (\mu_t - \mu_{t-1})^2/2 \right). \qquad (13)$$

5. Sample τ_θ from the following conjugate full-conditional:

$$\pi(\tau_\theta|\dots) \sim Gamma\left(a_\theta + N/2, b_\theta + \sum_i \theta_i^2/2 \right). \qquad (14)$$

6. Sample τ_γ from the following conjugate full-conditional:

$$\pi(\tau_\gamma|\dots) \sim Gamma\left(a_\gamma + \frac{N-1}{2}, b_\gamma + \sum_{j>1} \gamma_j^2/2 \right). \qquad (15)$$

In output, the algorithm returns a collection of sampled observations for each model parameter, which are then used to empirically characterize the targeted posterior distribution.

6 Model selection

In our analysis, we have 8 competing models, so we need to define an optimality criterion to select the best model for each dataset.

Some of the most commonly used criteria for model selection, especially in nested generalized regression models, are the AIC, or Akaike information criterion from Akaike (1974), and the BIC, or Bayesian information criterion from Schwarz (1978). In practice, these two criteria correspond to a penalised log-likelihood, where the penalization term grows with the number of parameters in the model considered, so that parsimony is promoted. These criteria define a quantitative framework to characterize the trade-off between model complexity and how good the model fit is.

In this paper, we adopt both the AIC and BIC as model choice criteria, and derive our optimal models using these approaches.

7 Simulation study

In this section, we illustrate our methodology on artificial data, to demonstrate its correctness and efficiency in estimating parameters and choosing the correct model. We consider artificial networks with $T = 10$ and $N = 40$, and model parameters generated as follows:

$$\mu^{(1)} \sim \mathcal{N}(0, \sigma_0^2); \qquad \mu^{(t+1)} \sim \mathcal{N}(\mu^{(t)}, \sigma_0^2); \qquad \theta_i \sim \mathcal{N}(0, \sigma_0^2); \qquad \gamma_j \sim \mathcal{N}(0, \sigma_0^2);$$

for all is, js and ts. In order to create an agreement with the constraint in Eq. (8), we apply a translation to $\gamma_1, \ldots, \gamma_N$ by subtracting $\sum_{j=1}^{N} \gamma_j / N$. We consider values of σ_0 in the set $\{0.25, 0.5, 1\}$, reflecting different levels of difficulty in estimating the model. This is due to the fact that the effect of each parameter diminishes with σ_0, since they all get closer to zero in our additive model. Vice versa, a larger σ_0 creates stronger effects which are easier to infer.

In addition, we generate data from four different models, as follows:

Model 1, `sender-receiver`:	$\log\left(\alpha_{ij}^{(t)}\right) = c + \theta_i + \gamma_j$
Model 2, `trend-receiver`:	$\log\left(\alpha_{ij}^{(t)}\right) = \mu^{(t)} + \gamma_j$
Model 3, `trend-sender`:	$\log\left(\alpha_{ij}^{(t)}\right) = \mu^{(t)} + \theta_i$
Model 4, `full`:	$\log\left(\alpha_{ij}^{(t)}\right) = \mu^{(t)} + \theta_i + \gamma_j$

Under each model, and under each value of σ_0, we generate 200 networks using our likelihood model, and then use our procedure to select the model and estimate all parameters. To do so, we run the full sampling procedure using the same four models as candidates, for each dataset. We use 1000 iterations as burn-in, and then save one sample every tenth, until a posterior sample of 100 values is obtained for each parameter. The proposal variances are all set to 1. Clearly, these sampling parameters are arbitrary, however, we show in the results that they work very well for the datasets considered.

Once we have run the sampling procedure for every σ_0, for every true model, for every candidate model, and for every network, we compare our candidate fitted models using the two model choice criteria: AIC and BIC.

The accuracy of the two model-choice criteria (expressed as the number of correctly identified models, out of 200) is shown in Table 2.

Here, we see that, generally, both criteria perform reasonably well under a variety of settings. This is a very positive result, considering that the real model parameters are generated with a fairly small standard deviation. In particular, AIC

Table 2 Confusion tables indicating the number of datasets (out of 200) for every combination of true model vs. estimated model, for three model selection criteria (over three rows of tables) and three values of σ_0 (over three columns of tables)

AIC					$\sigma_0 = 0.2$	AIC					$\sigma_0 = 0.3$	AIC					$\sigma_0 = 0.4$
		Estimated						Estimated						Estimated			
		1	2	3	4			1	2	3	4			1	2	3	4
True	1	200	0	0	0	True	1	200	0	0	0	True	1	200	0	0	0
	2	0	200	0	0		2	0	200	0	0		2	0	200	0	0
	3	0	0	200	0		3	0	0	200	0		3	0	0	200	0
	4	0	0	0	200		4	0	0	0	200		4	0	0	0	200
BIC					$\sigma_0 = 0.2$	BIC					$\sigma_0 = 0.3$	BIC					$\sigma_0 = 0.4$
		Estimated						Estimated						Estimated			
		1	2	3	4			1	2	3	4			1	2	3	4
True	1	153	47	0	0	True	1	200	0	0	0	True	1	200	0	0	0
	2	0	200	0	0		2	0	200	0	0		2	0	200	0	0
	3	0	0	200	0		3	0	0	200	0		3	0	0	200	0
	4	2	108	0	90		4	0	0	0	200		4	0	0	0	200

achieves the best results in recovering the correct model from the data, whereas BIC tends to infer a wrong model when $\sigma_0 = 2$. However, as expected, results generally improve when σ_0 increases. It is important to note that BIC tends to select Model 2 very often. This happens because Model 2 is the one containing the smallest number of parameters (T and $N - 1$, for $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$, respectively), so this is the model where the heavier penalization of the BIC has the least effect.

In Fig. 7, we show the estimation error for the best AIC models, for each of the model parameters and $\sigma_0 \in \{0.2, 0.4\}$.

This figure illustrates that, although the datasets are fairly small, the errors are very contained and there is little to no evidence of bias, regardless of the model considered. For the parameter $\boldsymbol{\theta}$, we notice some bias which is justifiable by the near-non-identifiability of the additive model (see Sect. 5.1). This bias seems to be compensated by the parameter $\boldsymbol{\mu}$. In both cases, points are still distributed on a line roughly parallel to the bisector line, indicating no errors up to a translations of the parameters.

8 OeNB 200 and OeNB 800 results

8.1 Technical details

We ran our Metropolis-within-Gibbs sampler on both datasets OeNB 200 and OeNB 800, and for all models, for a total of 100, 000 iterations. For all runs, the first 75, 000 iterations were discarded as burn-in. For the remaining sample, every
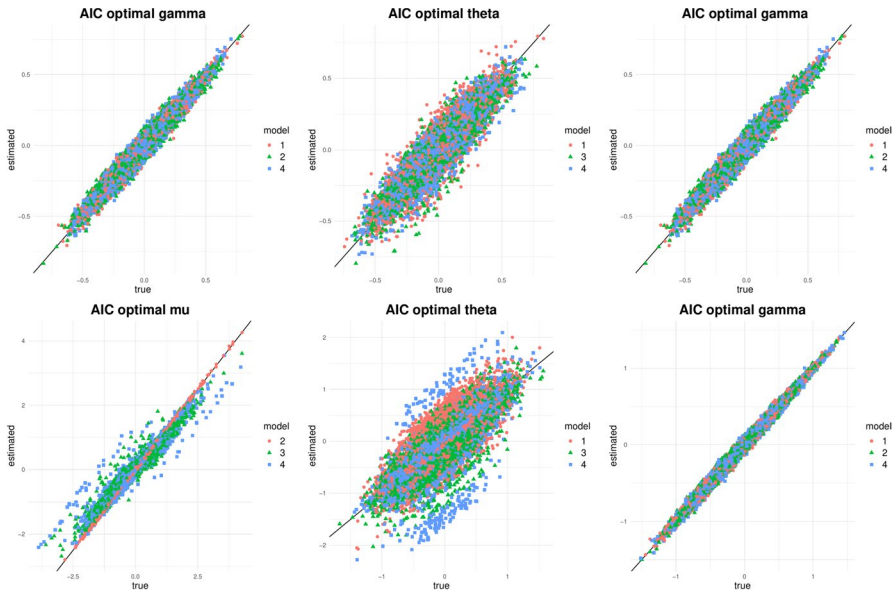
Fig. 7 True value of model parameters vs. estimated values. The top row corresponds to $\sigma_0 = 0.2$ and the bottom row corresponds to $\sigma_0 = 0.4$. Colors indicate the true underlying model number. The plots include only parameter estimates corresponding to correctly identified models

5-th draw was saved to produce the final results. In summary, we used samples of 5000 observations to characterize the posterior distribution of each model parameter.

The first 75, 000 iterations of the burn-in period were also used to adaptively tune the Gaussian proposal variance individually for each parameter, to make sure that all of the acceptance rates were approximately 33%. The proposal variances were hence fixed to these ideal values for the rest of the process. The trace plots and convergence diagnostic tests all showed very good mixing of the Markov chains, suggesting a satisfactory convergence.

Similarly to many other latent variable models for networks, the computational cost required by our sampler grows as TN^2. We implemented the algorithm in C++ to speed up the procedure, and then wrapped the functions into an R package which is available from our public repository (Rastelli and Hledik 2021).

8.2 Model comparisons

We calculate the AIC and BIC for all 8 models and for both datasets; the results are shown in Tables 3 and 4, which give more numerical details regarding the ranking of the models according to the criteria.

Table 3 `OeNB_200`: Ranking of the models according to AIC

Model	Parameters	Rescaled AIC	Rescaled BIC
sender-receiver	400	0.0	1413.6
full	415	33.2	1617.3
receiver	200	859.2	0.0
trend-receiver	215	896.4	207.6
sender	201	5677.7	4829.8
trend-sender	216	5731.7	5036.3
null	1	6190.7	3070.3
trend	16	6229.3	3279.1

The lowest AIC and BIC have been subtracted to all values in the respective columns

Table 4 `OeNB_800`: Ranking of the models according to AIC

Model	Parameters	Rescaled AIC	Rescaled BIC
receiver	800	0.0	2808.8
trend-receiver	815	44.1	3065.1
sender-receiver	1600	933.3	15054.6
full	1615	1039.6	15373.0
null	1	8489.5	0.0
trend	16	8533.5	256.2
sender	801	9601.2	12424.1
trend-sender	816	9643.4	12678.4

The lowest AIC and BIC have been subtracted to all values in the respective columns

We discuss the results according to the AIC and BIC optimal models in the following subsections.

8.3 `OeNB 200` **dataset**

In `OeNB 200`, the AIC-optimal model is `sender-receiver`, whereas the BIC-optimal model is `receiver`, which essentially corresponds to the same model where also θ has been removed. We further note that all models that include μ are worse than their simpler counterparts that do not include μ. This highlights very strong evidence that the global exposure diversification trend parameter μ is not relevant for this dataset, and thus it should not be included in the analysis.

In the context of our application, our procedure selects a model without the trend component, meaning that the exposure concentration level depends only on the bank's individual value and it does not depend on time. Thus, our method
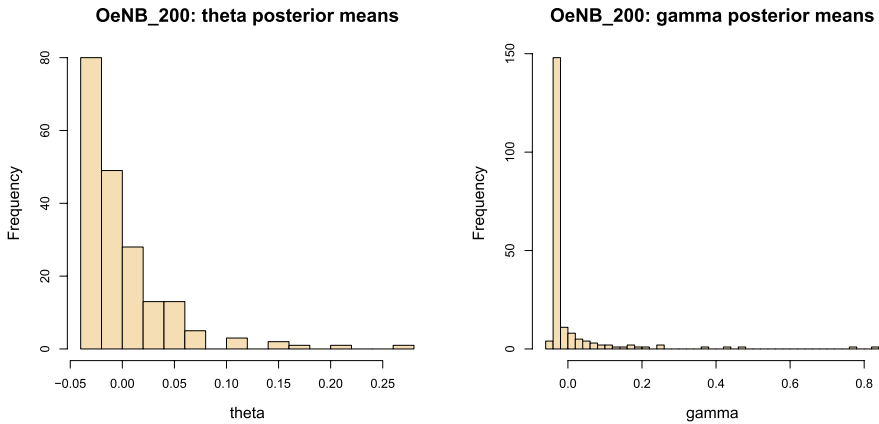
Fig. 8 Empirical distribution of posterior expectations of θ (left panel) and γ (right panel) for the 200 most relevant banks
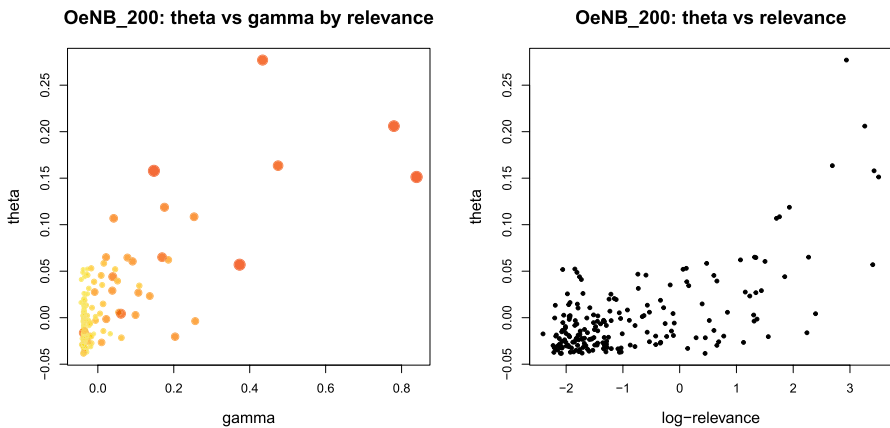

Fig. 9 Left panel: posterior expectations of θ vs. the posterior expectations of γ, for the `OeNB 200` dataset (`sender-receiver` model). Both size and color of the points indicate the relevance. Right panel: posterior expectations of θ vs. the aggregated relevance of banks, for `OeNB 200` (`sender-receiver` model)

indicates that the most relevant banks did not uniformly change their exposure concentration between 2018 and 2012, rather they adopted different strategies and reacted in different ways to the crises.

The AIC-optimal model `sender-receiver` characterizes the exposures of these banks using the attractiveness parameter of each bank (γ), and the individual exposure diversification of each bank (θ). We show the estimated posterior expectations of these parameters in Fig. 8. In both histograms we see that the distributions are skewed, in that there are some banks that have very large values of either exposure spread or attractiveness. In fact, from the left panel of Fig. 9, we see that those banks that have high exposure spread (low concentration) also tend
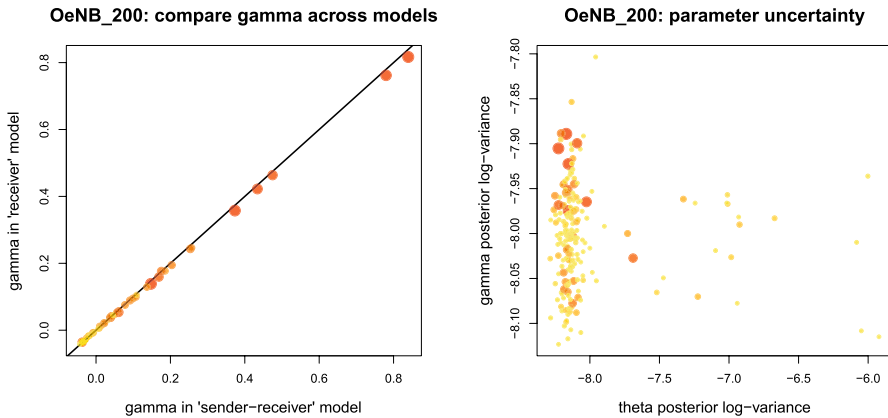
Fig. 10 Left panel: posterior expectations of γ in the `sender-receiver` model vs. posterior expectations of γ in the `receiver` model, for `OeNB 200`. Right panel: posterior variance of θ vs. posterior variance of γ, with nodes' sizes and colors representing aggregated relevance for `OeNB 200` (`sender-receiver` model)

to have high attractiveness, and vice versa. This figure highlights that large relevant banks tend to have higher exposure spread while simultaneously also being more attractive, and, vice versa, small banks often play a more peripheral role in the network, usually as offspring of some larger bank. A similar observation of heavy tailed degree distribution has also been reported by Boss et al. (2004).

The right panel of Fig. 9 focuses more directly on the relation between the estimates of θ and the aggregated relevance of banks. This scatter-plot indicates that more relevant banks tend to have a more pronounced exposure spread, whereas small banks concentrate their exposures more. This observation further confirms our ideas about a stylized financial network where the disassortative behavior is very common.

The left panel of Fig. 10 highlights instead a different aspect of our analysis: we show that the parameter estimates for γ are essentially the same for the two optimal models, indicating that the parameters θ and γ capture different aspects and patterns in the data.

This means that it is important to investigate the effect of both of these parameters through the model choice framework, since they give relevant and rather independent contributions in explaining the data. The reason why the θs are removed from one model is purely because of the penalization value which inevitably must lead to a model with fewer parameters.

The Bayesian approach employed for the inferential procedure permits a full characterization of the posterior distribution of each of the parameters. We take advantage of this to investigate whether there is any pattern in the quantified uncertainty around each parameter estimate. For this purpose, the right panel of Fig. 10 shows the relation between the posterior variances of the θ vs. γ. We note that there

seems to be no explicit pattern between the parameters, and no apparent relation with the relevance of the corresponding banks. The posterior variances of θs tend to be higher than the posterior variances of γs.

8.4 `OeNB 800` **dataset**

In `OeNB 800`, the model `receiver` is optimal according to the AIC, whereas the `null` model is optimal according to BIC. This is primarily determined by the fact that the full dataset involves a very large number of parameters, and thus the penalization term becomes substantial.

The `null` model is the simplest and it describes a homogeneous network structure, whereby all the banks exhibit the same behaviour at all time frames. While the expectation for all edge weights is the same, the variance associated to the Dirichlet distribution makes it possible that a variety of different values and patterns are observed. So, according to this model, it is still perfectly reasonable that some banks will be exposed to very few counterparts whereas others will distribute their risks on a more wide set of partners. Still, the most fundamental message that we extract from either of the optimal models is that there is no evidence of a relevant change over time, at a global level, for the connectivity patterns.

The optimal `receiver` model characterizes the exposures using only the attractiveness parameter for each bank. This is essentially equivalent to a fitness model, whereby banks with higher "fitness" are more often receiving edges than those with lower fitness. In our context, highly attractive banks tend to borrow from more partners, so they are key institutions to monitor since they can put pressure on a large number of other institutions.

In the left panel of Fig. 11, we see that the distribution of the γ posterior means is very heavy tailed, in that a number of institutions exhibit very high attractiveness values (histogram bins become very thin for values larger than 0.2). This highlights the heterogeneous structure of this network where the nodes can exhibit very diverse patterns of connections. The right panel of Fig. 11 underlines more clearly the connection to a fitness model, in that it essentially shows that the γ parameters clearly capture the in-degrees of the nodes. We also highlight the relation between γ and the relevance of the banks in the left panel of Fig. 12, where we notice that high relevance tends to translate into a high attractiveness.

For completeness, we show the posterior density for the only parameter c under the `null` model, in the right panel of Fig. 12.

9 Conclusions

This paper contributes to the literature on networks by extending the current available network frameworks to model temporal non-negatively weighted interactions, and to capture systematic parts of their development. Our original application to the Austrian interbank market gives a new perspective on the recent crises and demonstrates how our model can be used as a means to measure exposure concentration
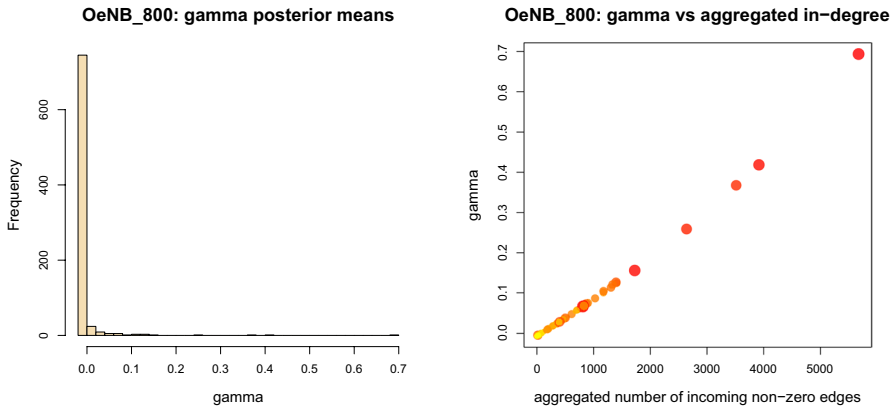
OeNB_800: gamma posterior means   **OeNB_800: gamma vs aggregated in−degree**

Fig. 11 Left panel: empirical distribution of the posterior means of γ under the `receiver` model. Right panel: empirical distribution of the posterior means of γ plotted against the total incoming weights for each node, under the `receiver` model. The color of points indicates the relevance
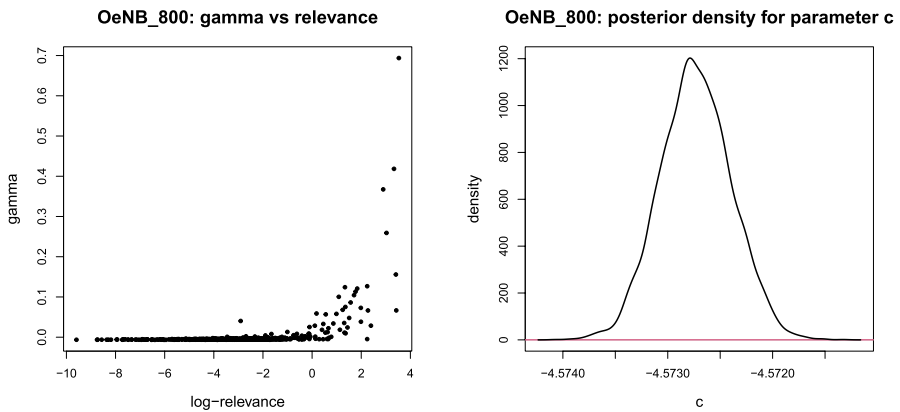
OeNB_800: gamma vs relevance   **OeNB_800: posterior density for parameter c**

Fig. 12 Left panel: posterior means of γ plotted against the log-relevance of each bank, for the `receiver` model. Right panel: empirical posterior density for the only parameter c in the `null` model

and bank's attractiveness as drivers of systemic risk. Differently from Friel et al. (2016), our measure is not affected by banks entering or leaving the system, since our available data only contains banks which are active throughout the whole period.

In our analysis we have shown that, for the Austrian market, the financial institutions did not change substantially their connectivity patterns over time, regardless of the financial instabilities experienced in 2008 and 2011. Our results are fairly in agreement with the exploratory data analysis, which highlighted some substantial changes with regard to the overall exposure sizes, but only one noticeable change in terms of concentration of exposures, during the second quarter of 2011. Our model-based approach showed that no change of patterns were significant, however we point out that our models that included the temporal dependency would capture a

major change in that same time period (results not shown). A critical reason why our approach does not consider the time trend significant is that the trend is assumed to be the same across all network participants. Clearly this can be a rather strong assumption, in that if time trends are present, then they would likely differ across participants. This strongly suggests a very interesting future development of this work whereby one can extend the model to include node-specific time trends.

The findings of our analysis may be of a particular use to regulators and central banks to assess and design future policy measures. In addition, our results showed that the roles played by the different banks can be vastly different, particularly in the context of exposure concentration. This emphasizes that larger banks, which are generally more susceptible to systemic risk, tend to use more conservative strategies and to spread out evenly their credit risks.

One limitation of our modeling framework is that it only focuses on the relative exposures, hence discarding the real magnitudes of the claims. Future extensions of this work may consider a joint modeling of the exposure values and how they are diversified among neighbors. Another possible extension of our framework would include a more sophisticated modeling structure, to account for transitivity or reciprocity. Also, one could easily resort to clustering, where, for example, different clusters are characterized by different network homogeneity drifts μ. Finally, we would like to remark that in our dataset all nodes are present at all time frames, however, this assumption is unlikely to hold in a more general setting. As a consequence, a relevant extension of our work would introduce mechanisms to let nodes join or exit the study at any point in time, perhaps in reflection of potential financial defaults.

Appendices

A. Data transformation

The source data from the Austrian National Bank is in the form of four variables: a timestamp, an ID of a lender bank, an ID of a borrower, and the *relative* exposure from one towards the other. We use the term *relative* since the largest exposure in *each* time period is assumed to be of size 1, and all other exposures *in that time period* are scaled accordingly to keep their relative size unchanged. As a result, in each time-period, all exposures are located in a (0, 1] interval with the highest exposure attaining a value of 1. Formally, making use of Definition 3.2, the observable data in our sample can be viewed as a dynamic adjacency matrix D:

Definition A.1 A *sequence of observable exposures* $\mathcal{D} = \{D^{(t)}\}_{t \in \mathcal{T}}$ on the set of nodes $\mathcal{V}$ over the time span $\mathcal{T}$ is defined as follows:

$$d_{ij}^{(t)} \overset{def}{=} \frac{e_{ij}^{(t)}}{\max_{k,l} e_{kl}^{(t)}} \quad \forall i,j,k,l \in \mathcal{V}, \forall t \in \mathcal{T} \tag{16}$$

It is not possible to make inter-temporal analysis of changes in exposures by using the sequence $\mathcal{D}$, because every exposure is scaled against the highest exposure in its time period. In order to circumvent this issue and obtain information which is comparable in time, we have devised the following procedure.

We make an assumption about the stability of the Austrian market. Namely, when looking at the change of a particular edge value between two consecutive periods, say from $d_{ij}^{(t)}$ to $d_{ij}^{(t+1)}$, the ratio $\frac{d_{ij}^{(t)}}{d_{ij}^{(t+1)}}$ with highest likelihood of occurrence in the sample corresponds to banks keeping the absolute value of their exposures unchanged. Indeed, after examining this ratio in all consecutive periods, we observe that the most frequent value is situated in the middle of the sample and is always a clear outlier in terms of likelihood of occurrence.[4]

It is straightforward to re-scale the whole dataset using this procedure. Despite the fact that we still cannot observe the actual levels of exposures between banks in our sample, we are now able to approximate these values up to a proportionality constant. We denote the values that we obtain with this procedure with $\mathcal{X}$ throughout the paper, and use these to calculate the relevance of the banks. We point out that these values are not used in the statistical model that we introduce.

To summarize, there are four different types of dynamic adjacency matrices used in our paper: $\mathcal{E}$ corresponds to the true data with the actual connection values (exact values not available), $\mathcal{D}$ represents the scaled data where edge weights are normalized with respect to the highest value in *each* period (exact values are available), $\mathcal{X}$ contains the scaled data where all edge weights are normalizes with respect to the highest value in the *first* period (available in approximate form), and $\mathcal{Y}$ contains the relative exposures of banks (exact values are available) which are derived from $\mathcal{X}$ or equivalently from $\mathcal{D}$.

---

[4] In most cases, this value is around 1 which suggests that the largest exposure in the network is mostly stable. An exception arises between dates 2 and 3 which correspond to the second and third quarter of 2008. As this is the exact time of the height of US subprime mortgage crisis, we believe that the "big players" in our dataset have been influenced by these events, resulting in the change of their exposures and subsequent substantial re-scaling of the whole system. According to our methodology, the largest exposure in the network has dropped to almost one third of its value in the span of two quarters, but it returns gradually back to its former level eventually.

References

Acemoglu D, Ozdaglar A, Tahbaz-Salehi A (2015) Systemic risk and stability in financial networks. Am Econ Rev 105:564–608

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19:716–723

Allen F, Gale D (2000) Financial contagion. J Polit Econ 108:1–33

Bardoscia M, Battiston S, Caccioli F, Caldarelli G (2017) Pathways towards instability in financial networks. Nat Commun 8:1–7

Battiston S, Martinez-Jaramillo S (2018) Financial networks and stress testing: challenges and new research avenues for systemic risk analysis and financial stability implications. J Financ Stab 35:6–16

Boss M, Elsinger H, Summer M, Thurner S (2004) Network topology of the interbank market. Quant Financ 4:677–684

Caccioli F, Barucca P, Kobayashi T (2018) Network models of financial systemic risk: a review. J Comput Soc Sci 1:81–114

Daudin J-J, Picard F, Robin S (2008) A mixture model for random graphs. Stat Comput 18:173–183

Elliott M, Golub B, Jackson MO (2014) Financial networks and contagion. Am Econ Rev 104:3115–53

Elsinger H, Lehar A, Summer M (2006) Risk assessment for banking systems. Manag Sci 52:1301–1314

Frey R, Hledik J (2018) Diversification and systemic risk: a financial network perspective. Risks 6:54

Friel N, Rastelli R, Wyse J, Raftery AE (2016) Interlocking directorates in Irish companies using a latent space model for bipartite networks. Proc Natl Acad Sci 113:6629–6634

Gai P, Kapadia S (2010) Contagion in financial networks. Proc R Soc Lond A Math Phys Eng Sci 466:2401–2423

Gai P, Kapadia S (2019) Networks and systemic risk in the financial system. Oxf Rev Econ Policy 35:586–613

Glasserman P, Young HP (2016) Contagion in financial networks. J Econ Lit 54:779–831

Goncharenko R, Hledik J, Pinto R (2018) The dark side of stress tests: negative effects of information disclosure. J Financ Stab 37:49–59

Hoff PD, Raftery AE, Handcock MS (2002) Latent space approaches to social network analysis. J Am Stat Assoc 97:1090–1098

Hurd TR (2016) Contagion! systemic risk in financial networks. Springer

Li M, Guan S, Wu C, Gong X, Li K, Wu J, Di Z, Lai CH (2014) From sparse to dense and from assortative to disassortative in online social networks. Sci Rep 4:4861

Matias C, Miele V (2017) Statistical clustering of temporal networks through a dynamic stochastic block model. J R Stat Soc Ser B (Stat Methodol) 79:1119–1141

Matias C, Rebafka T, Villers F (2018) A semiparametric extension of the stochastic block model for longitudinal networks. Biometrika 105:665–680

Minka T (2000) Estimating a dirichlet distribution. MIT

Nowicki K, Snijders TAB (2001) Estimation and prediction for stochastic blockstructures. J Am Stat Assoc 96:1077–1087

Puhr C, Seliger R, Sigmund M (2012) Contagiousness and vulnerability in the Austrian interbank market. Oesterreichische Nationalbank Financ Stab Rep 24:1–17

Rastelli R, Friel N, Raftery AE (2016) Properties of latent variable network models. Netw Sci 4:407–432

Rastelli R, Hledik J (2021) DirNet code repository. https://github.com/riccardorastelli/DirNet

Sarkar P, Moore AW (2006) Dynamic social network analysis using latent space models. Adv Neural Inf Process Syst 18:1145–1152

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464

Sewell DK (2018) Simultaneous and temporal autoregressive network models. Netw Sci 6:204–231

Sewell DK, Chen Y (2016) Latent space models for dynamic networks with weighted edges. Soc Netw 44:105–116

Upper C (2011) Simulation methods to assess the danger of contagion in interbank markets. J Financ Stab 7:111–125

van der Merwe S (2019) A method for Bayesian regression modelling of composition data. S Afr Stat J 53:55–64

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393:440

Westveld A H, Hoff P D (2011) A mixed effects model for longitudinal relational and network data, with applications to international trade and conflict

Xu KS, Hero AO (2014) Dynamic stochastic blockmodels for time-evolving social networks. IEEE J Sel Top Signal Process 8:552–562

Yang T, Chi Y, Zhu S, Gong Y, Jin R (2011) Detecting communities and their evolutions in dynamic social networks—a Bayesian approach. Mach Learn 82:157–189

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.