Working Paper Series

No 109 / April 2020

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

by
Juraj Hledik
Riccardo Rastelli

ESRB
European Systemic Risk Board
European System of Financial Supervision

ABSTRACT

We design a statistical model for measuring the homogeneity of a financial network that evolves over time. Our model focuses on the level of diversification of financial institutions; that is, whether they are more inclined to distribute their assets equally among partners, or if they rather concentrate their commitments towards a limited number of institutions. Crucially, a Markov property is introduced to capture time dependencies and to make our measures comparable across time. We apply the model on an original dataset of Austrian interbank exposures. The temporal span encompasses the onset and development of the financial crisis in 2008 as well as the beginnings of the European sovereign debt crisis in 2011. Our analysis highlights an overall increasing trend for network homogeneity, whereby core banks have a tendency to distribute their market exposures more equally across their partners.

Keywords: Latent Variable Models, Dynamic Networks, Austrian Interbank Market, Systemic Risk, Bayesian Inference

1 Introduction

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 re-adjust their portfolios in order to account for the change in the economic situation.

With the goal of contributing to the discussion on interbank exposures and diversification, 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 diversification levels of individual banks and of the market overall. We accomplish this by creating a new latent variable model for weighted networks that evolve over time. This framework provides us with a bank-dependent measure 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. 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.

It is important to understand that 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 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 Figure 1.1. In the case shown in Figure 1.1a, there is no danger of

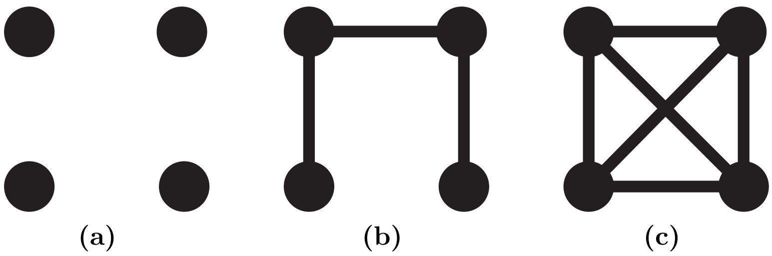


Figure 1.1. Different loan network structures on a set of four banks.

contagion since there are no edges to propagate shocks. An analogical result follows from the network shown in Figure 1.1c, 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 1.1b, 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 diversification 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 diversification of a financial system, hence obtaining a measure for one of the facets of systemic risk.

This paper is connected to two distinct strands of research. On the one 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] On the other hand, we also contribute towards theoretical papers dealing with latent variable modeling of network data. The method we propose borrows from and contributes to both fields, ultimately proposing a new perspective on systemic risk.

2 Related literature

One of the earliest papers on the topic of systemic risk in finance 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, Ozdaglar, and Tahbaz-Salehi (2015),

[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.

while Upper (2011) provides an excellent survey of regulatory-published scientific reports on the subject. With respect to the questions on exposure diversification, we refer the reader to Elliott, Golub, and Jackson (2014) and Frey and Hledik (2018), where a nontrivial relationship between diversification and contagious defaults is presented, or to Goncharenko, Hledik, and Pinto (2018), where banks endogenously choose their level of diversification in an equilibrium setting. 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 diversification and homogeneity of the system can be studied inter-temporally.

As mentioned, our paper also contributes to the research on latent variable network modeling. Prominent contributions include the latent position models of Hoff, Raftery, and Handcock (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, Chi, Zhu, Gong, and Jin (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, Friel, and Raftery (2016) and Daudin, Picard, and Robin (2008), 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, Rastelli, Wyse, and Raftery (2016), where the authors introduce a dynamic latent position model to measure the financial stability of the Irish Stock Exchange; Sewell and Chen (2016), who introduce a latent position model for dynamic weighted networks; and Matias, Rebafka, and Villers (2018), where the authors propose a dynamic extension of the stochastic blockmodel. Differently from these works, our approach relies on a new model whose goal is to measure and study the systemic risk associated to a financial system.

Lastly, we also contribute to the literature on the stability of the Austrian interbank market. Related works in this area include Elsinger, Lehar, and Summer (2006), Puhr, Seliger, and Sigmund (2012) and Boss, Elsinger, Summer, and Thurner (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 a an appreciation of the changes in the structure of the network over time.

3 Data and Exploratory Analysis

We use a unique 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. Nevertheless, for the purposes of our model, the true values of the claims are not required, since our approach 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 gather information on the importance of each institution. In order to better clarify these concepts, we now give a sequence of definitions to set our notation.

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 of 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 adjacency matrices of the same size $N \times N$, as in the following definition:

Definition 3.1. *A sequence of true exposures* $\mathcal{E} = \{\boldsymbol{E}^{(t)}\}_{t \in \mathcal{T}}$ *defined on the* *set of nodes $\mathcal{V}$ over the timespan $\mathcal{T}$ consists of a collection of adjacency matrices $\boldsymbol{E}^{(t)}$ $\forall t \in \mathcal{T}$ 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 $\mathcal{V} = \{1, \ldots, N\}$ and $\mathcal{T} = \{1, \ldots, T\}$. 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 due to privacy policy of the Austrian National Bank. For the purpose of this paper, we will therefore be working with the following quantities (see Appendix A for full details):

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

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

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. This normalization was a necessary requirement of the Austrian National Bank, yet it has no effect on the implications of our paper. The reason behind this transformation was to make the true euro amount transferred between institutions non-observable. However, from a modeling perspective, the monetary size is irrelevant and any normalization constant would be equally effective if applied to all exposures simultaneously. Therefore, we do not lose any information by not being able to observe the true exposures from 3.1 but only their scaled equivalents from 3.2 instead. Lastly, we define the sequence of relative exposures that our statistical model uses as observed data:

Definition 3.3. *A sequence of relative exposures* $\mathcal{Y} = \{\boldsymbol{Y}^{(t)}\}_{t\in\mathcal{T}}$ *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.

To summarize, $\mathcal{E}$ corresponds to the unobservable actual 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 only have a single use for $\mathcal{X}$, namely to create a subsample of "core banks". Essentially, selecting a portion of banks which can be deemed important allows us to see how the implications of our model are affected by the banks' size. In order to do so, we introduce the bank's *relevance*:

Definition 3.4. *A relevance of bank i in time period t is defined as:*

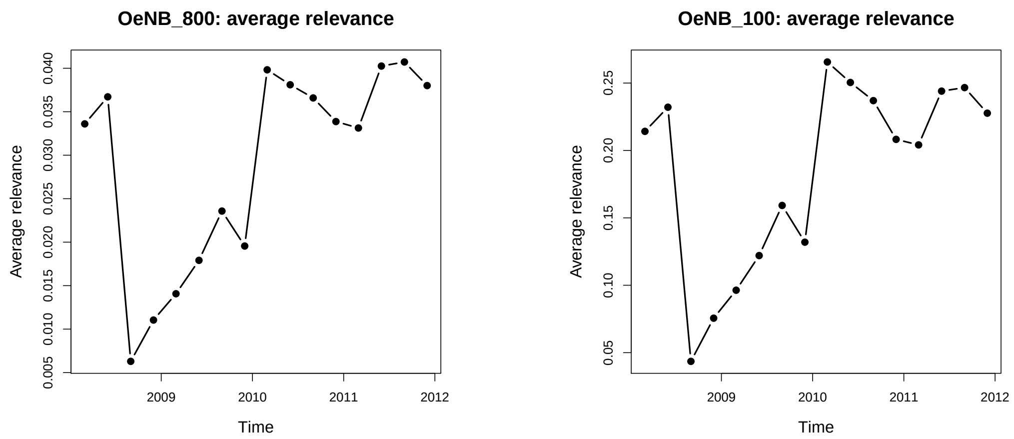
$$r_i^{(t)} = \sum_{k \in \mathcal{V}} x_{ik}^{(t)} + \sum_{k \in \mathcal{V}} x_{ki}^{(t)}. \tag{3}$$

In other words, we define relevance simply as the bank's overall sum of its interbank assets and liabilities.

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 100 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 100`, respectively.[2] We plot the evolution of the average bank relevance in Figure 3.1. As expected from its definition, we see a sharp drop in the second half of 2008 as a direct effect of the financial crisis.

In order to have a better picture about the data, we have conducted a brief exploratory analysis. Table 1 and Figure 3.2 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

[2]Validity of the `OeNB 100` subset can be justified further by examining the overall exposure of top 100 institutions. It turns out that the 100 most systemically relevant banks account for more than 95% of all approximate edge weights in any given time frame (according to the definition of absolute exposures from 3.2).



(a) OeNB 800 **(b)** OeNB 100

Figure 3.1. Bank relevance for the full sample (a) and the sample containing only the 100 most relevant banks (b).

column) depicts the overall cash flow in the market, scaled according to the first observation. We would like to highlight the second and third quarter of 2008, where a drop in the overall magnitude of cash flow 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 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 the first glance, there does not seem to be much in relation to this event in our data.

However, as we shall see later, our main model will provide further insight regarding the trend in diversification during this period.

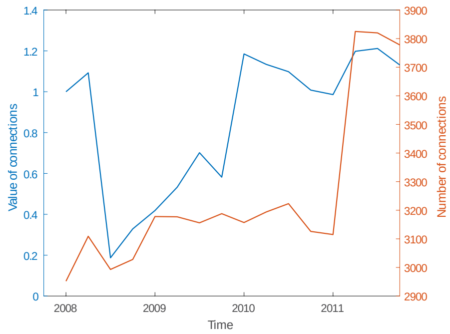


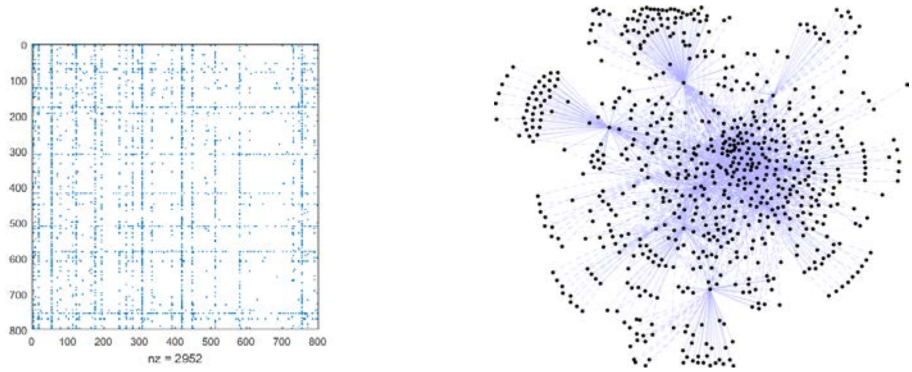
Figure 3.2. Number of connections and their relative size in time.

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 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, Guan, Wu, Gong, Li, Wu, Di, and Lai (2014); Watts and Strogatz (1998). Financial systems also tend to be very sparse. These same patterns are re-confirmed in the Austrian interbank market, as shown in Figure 3.3.

We have observed several interesting patterns in the data which suggest that using a more involved model could indeed produce new insights regarding the evolution of bank diversification. Since the main interest of our research lies in the diversification of agents in an interbank market, we have

Table 1. Number of connections and their relative size in time.

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



(a) Adjacency matrix **(b)** Plot of a network snapshot

Figure 3.3. Adjacency matrix for the first time period, 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).

also looked at the evolution of entropy in the 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:*

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

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

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.

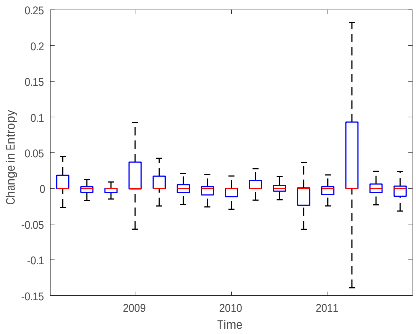


Figure 3.4. Distribution of entropy change in time.

In Figure 3.4, 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.

4 The 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):

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

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 $\alpha_{ij}^{(t)} > 0$. 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:

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

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 $\boldsymbol{\mu} = \left\{\mu^{(t)}\right\}_{t\in\mathcal{T}}$, $\boldsymbol{\theta} = \{\theta_i\}_{i\in\mathcal{V}}$ and $\boldsymbol{\gamma} = \{\gamma_j\}_{j\in\mathcal{V}}$ possess a straightforward interpretation, which we illustrate in the next section.

We point out that our model can be thought of as a Dirichlet regression model for compositional data, see Minka (2000), van der Merwe (2019) and references therein.

4.1 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 $\mathbf{y} \sim Dir\left(\boldsymbol{\alpha}\right)$. If the value of α increases, then the variance of 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. 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 Section 3, respectively.

In our formulation, a similar reasoning holds, even if the $\boldsymbol{\alpha}$ vector is not symmetric. The log-additive structure in (7) deterministically 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, we are essentially capturing the level of homogeneity in the network through a homogeneity trend parameter $\mu^{(t)}$ and a node specific homogeneity random effect θ_i. In other words, an increase in either $\mu^{(t)}$ or θ_i corresponds to higher diversification 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 a decrease in diversification which in turn results in a more heterogeneous network structure.

The interpretation of γ_j is similar. To see this, consider a non-symmetric random vector $\mathbf{y} \sim Dir(\alpha_1, \ldots, \alpha_N)$. In this case, 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 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. Parameter $\mu^{(t)}$ indicates the global homogeneity level at time frame $t \in \mathcal{T}$, parameter θ_i characterizes the individual bank i homogeneity level as a random effect, and parameter γ_j represents the bank j's attractiveness.

4.2 Bayesian hierarchical structure

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

We assume a random walk process prior on the drift parameters $\boldsymbol{\mu}$ as follows:

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

where $\eta_t \sim \mathcal{N}(0, 1/\tau_\eta)$ and $\tau_\eta \sim Gamma(a_\eta, b_\eta)$. The 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.

The parameters $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ are assumed to be i.i.d. Gaussian variables with:

$$\begin{aligned} \theta_i &\sim \mathcal{N}\left(0, 1/\tau_\theta\right), & \tau_\theta &\sim \text{Gamma}(a_\theta, b_\theta), \\ \gamma_j &\sim \mathcal{N}\left(0, 1/\tau_\gamma\right), & \tau_\gamma &\sim \text{Gamma}(a_\gamma, b_\gamma) \end{aligned}$$

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 Figure 4.1 summarizes the dependencies in our model graphically.

5 Parameter estimation

Our proposed model has T drift parameters (denoted by $\boldsymbol{\mu}$), N 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.

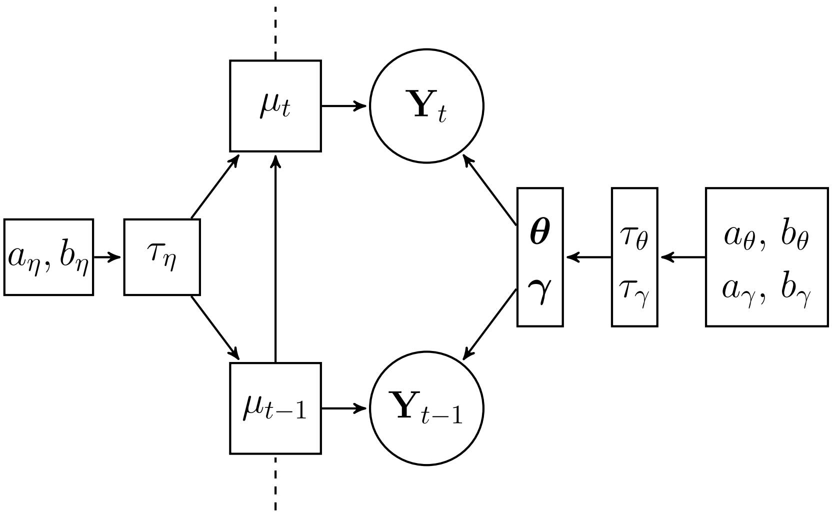


Figure 4.1. Graphical representation of model dependencies.

5.1 Identifiability

The additive structure in (7) yields a non-identifiable likelihood model. For example, one could define $\tilde{\theta}_i = \theta_i + c$ and $\tilde{\gamma}_j = \gamma_j - c$ for some $c \in \mathbb{R}$ and the likelihood value would be the same for the two configurations, i.e. $\mathcal{L}_{\mathcal{Y}}\left(\boldsymbol{\mu}, \tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\gamma}}\right) = \mathcal{L}_{\mathcal{Y}}\left(\boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\gamma}\right)$. 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 accepted 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 $T + 2N + 2$ parameters, is now identifiable.

5.2 Markov chain Monte Carlo

The posterior distribution associated to our model factorizes as follows:

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

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

[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.

Metropolis-Hastings with a Gaussian proposal:

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

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

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

3. Sample γ_ℓ for all $\ell \in \mathcal{V} \setminus \{1\}$ from the following full-conditional using

Metropolis-Hastings with a Gaussian proposal:

$$\pi\left(\gamma_\ell|\ldots\right) \propto \left\{\prod_{t,i}\Gamma\left(e^{\mu_t}e^{\theta_i}\sum_j e^{\gamma_j}\right)\right\}\left\{\prod_{t,i}\frac{\left[y_{i\ell}^{(t)}\right]^{\alpha_{i\ell}^{(t)}-1}}{\Gamma\left(\alpha_{i\ell}^{(t)}\right)}\right\} \cdot \left\{\prod_{t,i}\frac{\left[y_{i1}^{(t)}\right]^{\alpha_{i1}^{(t)}-1}}{\Gamma\left(\alpha_{i1}^{(t)}\right)}\right\}\exp\left\{-\frac{\tau_\gamma}{2}\gamma_\ell^2\right\}. \tag{12}$$

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

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

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

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

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

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

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.

5.3 Technical details

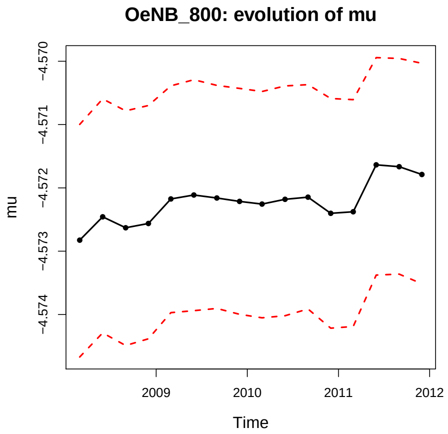
We ran our Metropolis-within-Gibbs sampler on both datasets `OeNB 800` and `OeNB 100` for a total of 400,000 iterations. For both datasets, the first 200,000 iterations were discarded as burn-in. For the remaining sample, every 20-th draw was saved to produce the final results. In summary, we used samples of 10,000 observations to characterize the posterior distribution of each model parameter.

The first 100,000 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 between 22% and 30%. The variances were hence fixed to these values for the rest of the process. The trace plots and convergence diagnostic tests all showed very good mixing of the Markov chain, suggesting a satisfactory convergence.

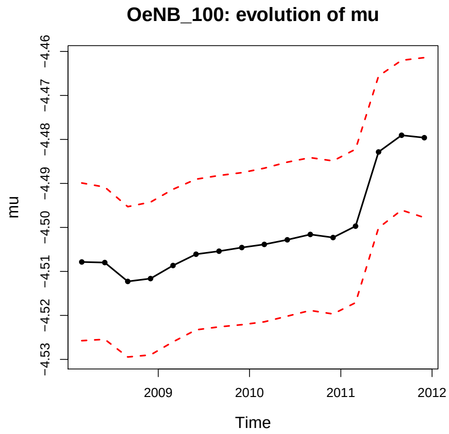
Similarly to many other latent variable models for networks, the computational cost required by our sampler grows as TN^2. We implemented the algorithm in `C++` and used parallel computing via the library `OpenMPI` to speed up the procedure. We note that, for the full dataset, an iteration required an average of approximately 0.75 seconds on a Debian machine with 16 cores. The code is available from the authors upon request.

6 Results

First, we study the diversification of the banks which translates to changes in network homogeneity. The drift parameter $\mu^{(t)}$, shown in Figure 6.1, exhibits an upward trend for both datasets. This trend is in both cases more



(a) OeNB 800



(b) OeNB 100

Figure 6.1. Evolution of the posterior mean of $\mu^{(t)}$ for the full sample (a) and the sample containing only the 100 most relevant banks (b), with 95% credible intervals.

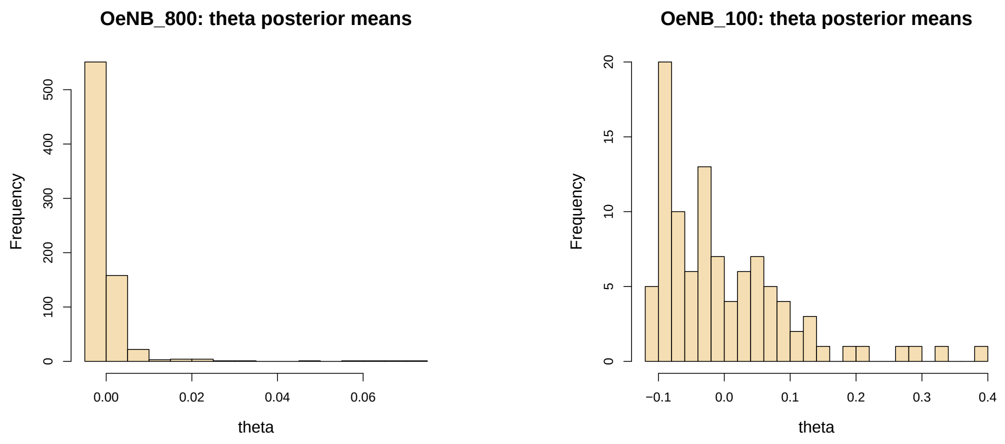
pronounced during the onset of the 2011 sovereign debt crisis. Furthermore, we observe a sharper increase in `OeNB 100` during this time period. This signals that larger and systemically relevant banks were the ones with a stronger reaction to the crisis. This shows not only their change in diversification policy as a reaction to the crisis, but also a difference in risk aversion for the two classes of agents in the network. Interestingly, we do not observe a similar behavior during the crisis in 2008.

In the exploratory analysis conducted earlier, we have seen a substantial drop in the overall exposure size in 2008 and almost no such effect in 2011. Paradoxically, we observe a large upward shift in diversification in 2011, while the same effect in 2008 is limited at best. One takeaway from this would be that Austrian banks have perceived the sovereign crisis as a bigger threat than the 2008 crisis stemming from the US housing market. As the effect is more pronounced in the OeNB 100 sample, it hints at the fact that bigger banks tend to increase their level of diversification more substantially, while less relevant banks tend to keep their exposures less diversified.

In addition to the overall development of diversification in the system, we can also study the same index locally to outline the aversion of individual banks to risk. This can be achieved by observing the parameter θ_i which denotes the diversification random effect value.

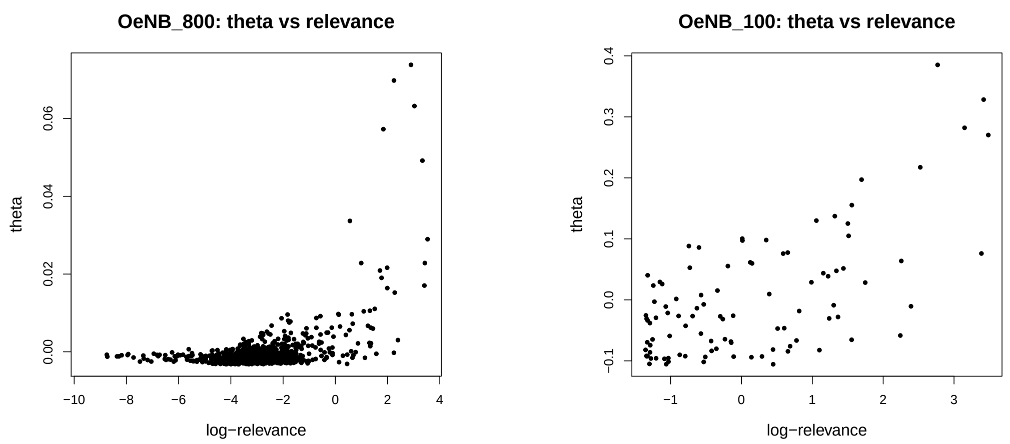
First, we analyze point estimates of these parameters: Figure 6.2 shows the distribution of the posterior means of $\boldsymbol{\theta}$. For both `OeNB 800` as well as `OeNB 100`, the distribution seems to be rather heavy tailed. This translates to a system where the majority of banks exhibits low diversification, but still a fairly large number of banks tends to diversify much more. In fact, Figure 6.3 highlights that more relevant banks tend to have a more pronounced diversification, whereas small banks do not diversify as much. This observation further confirms our ideas about a stylized financial network where the disassortative behavior is very common.

A similarly heavy tailed distribution can be observed regarding the at-



(a) OeNB 800 **(b)** OeNB 100

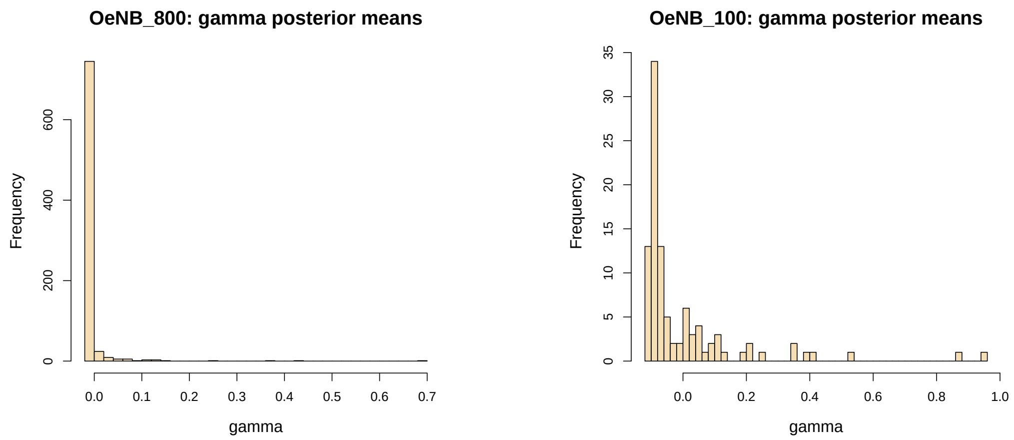
Figure 6.2. Posterior distribution of $\boldsymbol{\theta}$ for the full sample (a) and the sample containing only the 100 most relevant banks (b).



(a) OeNB 800 **(b)** OeNB 100

Figure 6.3. The banks with higher aggregated relevance tend to also have a higher diversification of exposures in both datasets.

tractiveness parameter $\boldsymbol{\gamma}$ (see Figure 6.4 for the distribution of the point estimates, where the heavy right tail is apparent). In addition, Figure 6.5

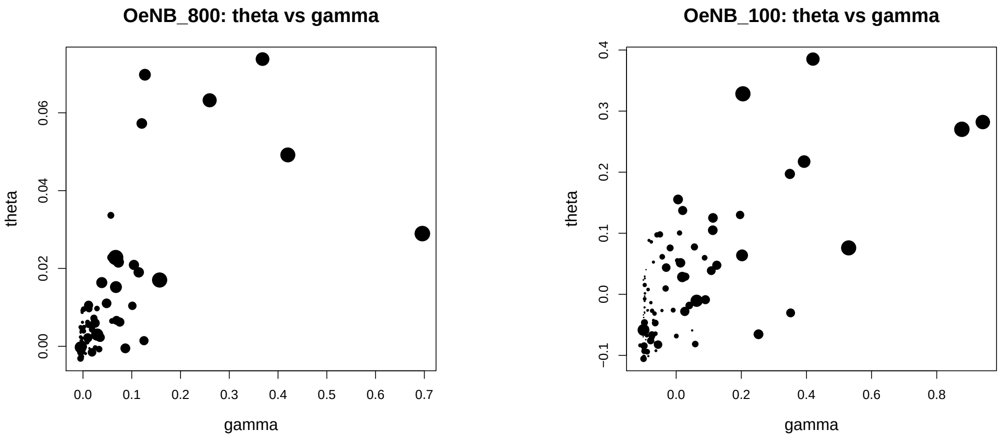


(a) OeNB 800 **(b)** OeNB 100

Figure 6.4. Posterior distribution of $\boldsymbol{\gamma}$ for the full sample (a) and the sample containing only the 100 most relevant banks (b).

shows that, generally, $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ are closely related in both datasets. This figure highlights that larger banks tend to be more diversified and more attractive simultaneously, and, vice versa, small banks often play a more peripheral role in the network, usually as offsprings of some larger bank. A similar observation of heavy-tailedness in degree distribution has also been reported by Boss et al. (2004).

As concerns the uncertainty around the point estimates, Figure 6.6 compares the posterior variances for all of the θs with those of the γs. We note that there seems to be no explicit pattern and no apparent relation with the relevance of the corresponding banks. We point out, however, that the two



(a) OeNB 800 **(b)** OeNB 100

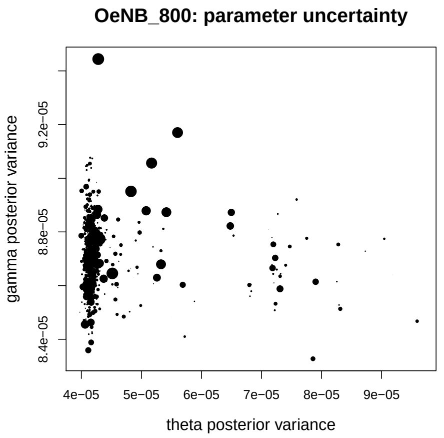
Figure 6.5. Posterior distribution of γ for the full sample (a) and the sample containing only the 100 most relevant banks (b). In both plots, the size of each circle represents the aggregated relevance of the corresponding bank.

plots are on two different scales on both axes, which is expected since much more data is available for inference in the `OeNB 800` dataset, hence yielding more reliable estimates.

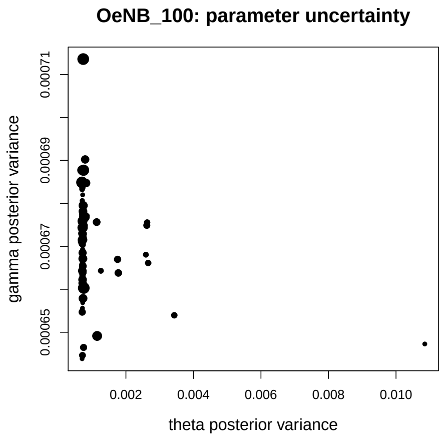
Finally, we also show the posterior densities for the variance parameters $1/\tau_\eta, 1/\tau_\theta$ and $1/\tau_\gamma$ in Figure 6.7. For both datasets, these plots suggest that the drift parameter is rather stable over time, and that the diversification and attractiveness are not particularly diverse across banks, overall.

7 Conclusion

This paper contributes to the literature on networks by proposing a brand new framework to model the evolution of dynamic weighted interactions and



(a) OeNB 800

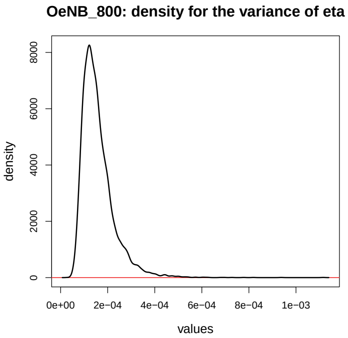


(b) OeNB 100

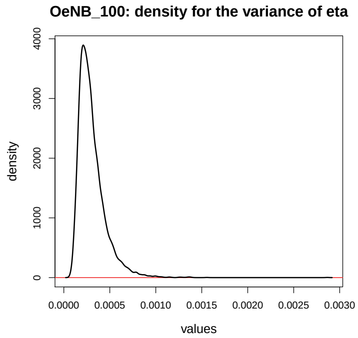
Figure 6.6. Posterior variances of $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ for the full sample (a) and the sample containing only the 100 most relevant banks (b). In both plots, the size of each circle represents the aggregated relevance of the corresponding bank.

to capture systematic parts of their development. Our 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 diversification as one of the components of systemic risk. Differently from Friel et al. (2016), our measure is not affected by banks entering or leaving the system, since our dataset only contains banks which are active throughout the whole period.

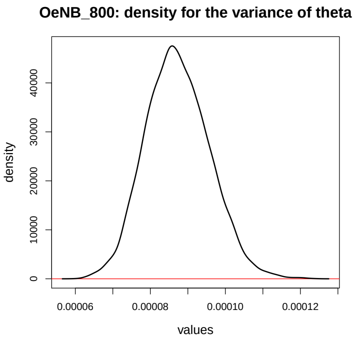
In our analysis we have shown that the Austrian market exhibited a sustained increase in banks' diversification, possibly as a reaction to the 2008 financial crisis. In particular, differently from a simple descriptive analysis, our model was able to capture a distinct upward dynamic in network homo-



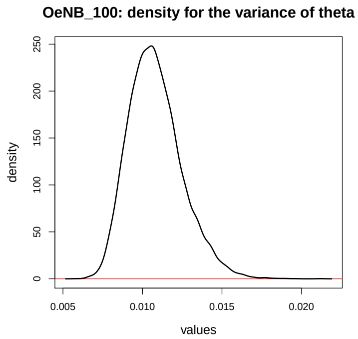
(a) $1/\tau_\eta$, `OeNB 800`



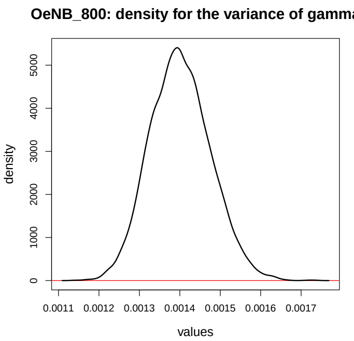
(b) $1/\tau_\eta$, `OeNB 100`



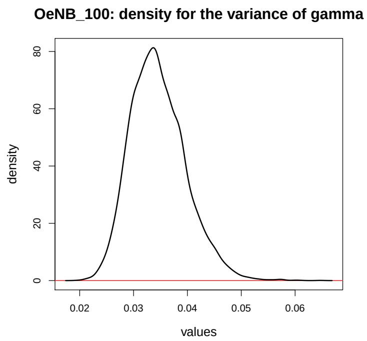
(c) $1/\tau_\theta$, `OeNB 800`



(d) $1/\tau_\theta$, `OeNB 100`



(e) $1/\tau_\gamma$, `OeNB 800`



(f) $1/\tau_\gamma$, `OeNB 100`

Figure 6.7. Posterior distribution of variance parameters $1/\tau_\eta, 1/\tau_\theta$ and $1/\tau_\gamma$ for `OeNB 800` (left) and `OeNB 100` (right) datasets.

geneity as a response to the sovereign debt crisis of 2011. These findings may be of a particular use to regulators and central banks to assess and design future policy measures.

Our results also showed that the roles played by the different banks can be vastly different, particularly in the context of exposure diversification. Our findings emphasize 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 prior structure on the model parameters. With further information on bank fundamentals, one could easily resort to clustering, where different clusters are characterized by different network homogeneity drifts $\boldsymbol{\mu}$.

Finally, we would like to remark that there are potentially other possible specifications besides the Dirichlet likelihood. The flexibility of Dirichlet distribution is known to be rather limited, as - for high levels of variance - it tends to assign most of the probability density to the highest entropy configurations. This does not necessarily reflect the features exhibited by the data. However, we argue that in our application the Dirichlet assumption is very reasonable, and, more importantly, it provides a convenient framework with a straightforward interpretation of the model parameters.

References

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

Allen, F., and D. Gale, 2000, Financial contagion, *Journal of political economy* 108, 1–33.

Boss, M., H. Elsinger, M. Summer, and S. Thurner, 2004, Network topology of the interbank market, *Quantitative finance* 4, 677–684.

Daudin, J.-J., F. Picard, and S. Robin, 2008, A mixture model for random graphs, *Statistics and computing* 18, 173–183.

Elliott, M., B. Golub, and M. O. Jackson, 2014, Financial networks and contagion, *American Economic Review* 104, 3115–53.

Elsinger, H., A. Lehar, and M. Summer, 2006, Risk assessment for banking systems, *Management science* 52, 1301–1314.

Frey, R., and J. Hledik, 2018, Diversification and systemic risk: A financial network perspective, *Risks* 6.

Friel, N., R. Rastelli, J. Wyse, and A. E Raftery, 2016, Interlocking directorates in irish companies using a latent space model for bipartite networks, *Proceedings of the National Academy of Sciences* 113, 6629–6634.

Gai, P., and S. Kapadia, 2010, Contagion in financial networks, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, rspa20090410, The Royal Society.

Glasserman, P., and H. P. Young, 2016, Contagion in financial networks, *Journal of Economic Literature* 54, 779–831.

Goncharenko, Roman, Juraj Hledik, and Roberto Pinto, 2018, The dark side of stress tests: Negative effects of information disclosure, *Journal of Financial Stability* 37, 49–59.

Hoff, P. D., A. E. Raftery, and M. S. Handcock, 2002, Latent space approaches to social network analysis, *Journal of the American Statistical Association* 97, 1090–1098.

Hurd, T. R., 2016, *Contagion! Systemic Risk in Financial Networks* (Springer).

Li, M., S. Guan, C. Wu, X. Gong, K. Li, J. Wu, Z. Di, and C. H. Lai, 2014, From sparse to dense and from assortative to disassortative in online social networks, *Scientific reports* 4, 4861.

Matias, C., and V. Miele, 2017, Statistical clustering of temporal networks through a dynamic stochastic block model, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79, 1119–1141.

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

Minka, T., 2000, Estimating a dirichlet distribution.

Nowicki, K., and T. A. B. Snijders, 2001, Estimation and prediction for stochastic blockstructures, *Journal of the American Statistical Association* 96, 1077–1087.

Puhr, C., R. Seliger, and M. Sigmund, 2012, Contagiousness and vulnerability in the austrian interbank market, *Oesterreichische Nationalbank Financial Stability Report* 24.

Rastelli, R., N. Friel, and A. E. Raftery, 2016, Properties of latent variable network models, *Network Science* 4, 407–432.

Sarkar, P., and A. W. Moore, 2006, Dynamic social network analysis using latent space models, in *Advances in Neural Information Processing Systems*, 1145–1152.

Sewell, D. K., and Y. Chen, 2016, Latent space models for dynamic networks with weighted edges, *Social Networks* 44, 105–116.

Upper, C., 2011, Simulation methods to assess the danger of contagion in interbank markets, *Journal of Financial Stability* 7, 111–125.

van der Merwe, S., 2019, A method for bayesian regression modelling of composition data, *South African Statistical Journal* 53, 55–64.

Watts, D. J., and S. H. Strogatz, 1998, Collective dynamics of ‘small-world’networks, *nature* 393, 440.

Xu, K. S., and A. O. Hero, 2014, Dynamic stochastic blockmodels for time-evolving social networks, *IEEE Journal of Selected Topics in Signal Processing* 8, 552–562.

Yang, T., Y. Chi, S. Zhu, Y. Gong, and R. Jin, 2011, Detecting communities and their evolutions in dynamic social networks – a bayesian approach, *Machine learning* 82, 157–189.

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:*

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

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]

[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 rescaling 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.

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 $\mathcal{D}$.

Imprint and acknowledgements

The authors developed this paper while affiliated with the WU Vienna University of Economics and Business, Austria. R. R. is currently affiliated with the School of Mathematics and Statistics, University College Dublin, Ireland. The authors kindly acknowledge the financial support of the Austrian Science Fund (FWF) as well as the possibility to use data provided by Austrian National Bank (OeNB). This research was also supported by the Vienna Science and Technology Fund (WWTF) Project MA14-031. All errors are the authors' sole responsibility.

Juraj Hledik
European Central Bank, Frankfurt am Main, Germany; email: **juraj.hledik@ecb.europa.eu**

Riccardo Rastelli
University College Dublin, Dublin, Ireland; email: **riccardo.rastelli@ucd.ie**

Postal address	60640 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Website	**www.esrb.europa.eu**

Note:
The views expressed in ESRB Working Papers are those of the authors and do not necessarily reflect the official stance of the ESRB, its member institutions, or the institutions to which the authors are affiliated.

ISSN	2467-0677 (pdf)
ISBN	978-92-9472-135-8 (pdf)
DOI	10.2849/150699 (pdf)
EU catalogue No	DT-AD-20-003-EN-N (pdf)