



Department of Economics and Management

DEM Working Paper Series

**Graphical network models for international
financial flows**

Paolo Giudici
(Università di Pavia)

Alessandro Spelta
(Università di Pavia)

52 (10-13)

Via San Felice, 5
I-27100 Pavia
<http://epmq.unipv.eu/site/home.html>

October 2013

Graphical network models for international financial flows

P. Giudici¹

Department of Economics and Management, University of Pavia

A. Spelta²

Department of Economics and Management, University of Pavia

Abstract

The late-2000s financial crisis has stressed the need of understanding the world financial system as a network of countries, where cross-border financial linkages play a fundamental role in the spread of systemic risks. Financial network models, that take into account the complex interrelationships between countries, seem to be an appropriate tool in this context. In this paper we propose to enrich the topological perspective of network models with a more structured statistical framework, that of graphical Gaussian models, which can be employed to accurately estimate the adjacency matrix, the main input for the estimation of the interconnections between different countries. We consider different types of graphical models: besides classical ones, we introduce Bayesian graphical models, that can take model uncertainty into account, and dynamic Bayesian graphical models, that provide a convenient framework to model temporal cross-border data, decomposing the model into autoregressive and contemporaneous networks. The paper shows how the application of the proposed models to the Bank of International Settlements locational banking statistics allows the identification of four distinct groups of countries, that can be considered central in systemic risk contagion.

Key words: Financial network models, Graphical models, Bayesian model selection

1 Introduction

Globalization of the economies leads to an ever-increasing interdependence of countries. Despite the conventional wisdom that national economies are

¹ paolo.giudici@unipv.it

² alessandro.spelta01@universitadipavia.it

becoming more and more interdependent, there is less evidence on the actual contribution of economic crises to the reinforcement of cross-border linkages and, in particular, of financial ones. The late-2000s financial crisis - considered by many economists to be the worst financial crisis since the Great Depression - resulted in the collapse of large financial institutions, in the bailout of banks by national governments and in downturns in stock markets around the world. Such a large set of outcomes has stressed the need of understanding financial systems as networks of countries, where cross-border financial linkages play a fundamental role.

Deriving macro (global) situations from micro (local) scenarios has been a recurrent topic in economics. A way to link the macro and the micro levels hinges on network theory, which has been introduced in economics to study formally the generation and the stability of economic interactions among agents (Watts and Strogatz, 1998, Jackson and Watts, 2002, Billio et al., 2012). Even though basic network structures have been introduced long time ago by Helpman and Krugman (1985), to explain international trade relations, the research knowledge on financial networks is far from being sufficient to exploit the potential of this powerful topological approach.

The seminal contributions in this field have focused on the study of interdependencies between financial systems, as in Kubelec and Sa (2012), particularly under shocks storming the financial systems of several countries altogether, as in Helpman and Krugman (1985). In particular, the topological properties of national interbank markets have been studied by Soramaky and co-authors (2007), who analyzed the network topology of interbank payments transferred between commercial banks by the Fedwire Funds Services. A related contribution is the work of Fuijwara et al. (2009), that explores the credit relationships that exist between commercial banks and large companies in Japan. Empirical network studies have also been carried out on some European national interbank markets (De Masi et al., Boss et al., 2004).

While the previous contributions focus mainly on local interbank markets, other papers have addressed the evolution of networks of bank transfers at a more global level, using the Bank of International Settlements (BIS) data set (Garratt et al., 2011, McGuire and Tarashev, 2006, Minoiu and Reyes, 2011). In particular, the work of McGuire and Tarashev (2006) shows that the international banking system has become an important conduit for the transfer of capitals across countries, whereas Minoiu and Reyes (2011) found evidence of important structural changes in banking financial networks, following the occurrence of stress events.

Our aim here is to further pursue the previous lines of research, using a stochastically structured financial network modeling approach to understand the impact of financial crises on the international banking system and, in par-

ticular, to understand which countries are most central and, therefore, most contagious (or subject to contagion).

As recently argued in Spelta and Araujo (2012), Billio et al (2012) and Barigozzi and Brownlees (2013), the adoption of a robust financial network approach in the above context is recommended not only because of its proper emphasis on financial interdependencies, but also for its possibility to describe how the structure of these interdependencies evolves in time. Were this be achieved, we would be able to address systemic financial risks in two directions: on one hand, to understand the role that a cross-border network structure plays in the spread of financial shocks; on the other hand, to understand the impact of stress events on a network of cross-border interdependencies.

Related to our work are recently proposed network models, based on simulated and real data (Nier et al., 2008, Gai and Kapadia, 2009, Georg and Poschmann, 2010, Castren and Kavonius, 2009, Markose et al., 2010, Cont et al., 2010, Peltonen et al., 2013). Simulated networks can help studying the impact of crisis scenarios on the system, in a stress-testing framework. However, they typically assume a fixed network structure and, therefore, are not suited for our purposes: learning financial interconnectedness from the data.

Our proposal will instead be based on real Bank of International Settlements data, as in Garratt et al (2011), McGuire and Tarashev (2006), Minoiu and Reyes (2011). Our additional contribution will be the proposal of a novel financial network model, based on stochastic graphical models, that enriches the network perspective with a more structured statistical approach. A more structured approach allows results that are more robust with respect to data variations and, in addition, being a full inferential approach, properly adjusts statistical estimates taking sample variability into account. These advantage are particularly evident when statistical inferences concern predictions of future events, such as contagion from actual stress.

Graphical models can be informally defined as a “marriage” between multivariate statistics and graph theory. They embed the idea that interactions among random variables in a system can be represented in the form of graphs, whose nodes represents the variables and whose edges shows their interactions. For an introduction to graphical models see, for instance, Pearl (1988), Lauritzen and Wermuth (1989), Whittaker (1990), Wermuth and Lauritzen (1990), Edwards (1990), Lauritzen (1996).

In this contribution we suggest that graphical models can be employed to accurately estimate the adjacency matrix, aimed at measuring interconnectedness between the different agents and, in particular, to assess central ones (as in Billio et al., 2012).

Traditional network models use the correlation matrix estimated from the

data to derive the adjacency matrix. Although useful, this approach takes into account only the marginal (indirect) effect of a variable on another, without looking at the (direct) effect of other variables. In our context, it does not distinguish between the direct and the indirect effect of a country on other ones.

Graphical models, instead, focus on the partial correlation matrix, that is obtained by measuring only the direct correlation between two variables. A partial correlation coefficient can express the change in the expected value of a dependent variable, caused by a unitary change of the independent variables, when the remaining variables are held constant. In so doing, the effect of a country on another is split into a direct effect (estimated by the partial correlation) and an indirect effect (what is left in the marginal correlation). Here we follow this approach and derive the adjacency matrix, the main input of a financial network model, not from the correlation matrix but, rather, from the partial correlation matrix obtained from the application of graphical models to the available data.

To achieve this aim we consider multivariate Gaussian graphical models, defined in terms of Markov properties, that is, conditional independencies associated with the underlying graph (Lauritzen, 1996; Whittaker, 1990). While traditional network models assume fixed graphical structures (such as fully connected graphs), the structure of a graphical model is typically random, and can be learned from the data, as a good fitting structure. Such a model selection can be performed by testing, in a stepwise procedure, the statistical significance of conditional independencies, which are equivalent to specified zeroes among certain partial correlation coefficients which, in turn, are equivalent to missing edges in the network representation.

The use of graphical models can thus help to have a deeper understanding of the relationships between variables, by distinguishing direct from indirect relationships. From their appearance in the 90's, several methodological advances have been made for graphical models. Less so in terms of applications, especially in economics. However, the application of graphical models to other fields (especially in the biomedical sciences) have highlighted two main problems, that require more advanced formulations.

First, the assumption of a random graph may be an important added value in situations where little a priori knowledge is present, as in several biological, machine learning and data mining applications. On the other hand, in economics a graphical structure must provide not only a good fit but also a good interpretation. This may result in choosing a model that has little support from the data, leading to predictions worse than could be obtained with other models. In other words inference restricted to be “model-dependent” may lead to suboptimal results. A second problem is that graphical models are essen-

tially static, photographing a situation in a given time span. This assumption seems to be restrictive in economics, in the case of variables that change over time, for example during periods of financial stress.

The above problems can be solved with the use of more advanced, Bayesian, graphical models, as shown in Madigan and York (1995), Giudici and Green (1999) and, more recently, Ahelegbey et al. (2012). In particular, Madigan and York (1995) and Giudici and Green (1999) propose a Bayesian model able to consider all possible graphical structures, choose the best fitting ones and, if necessary, average inferential results over the set of all models, thus solving the first problem. In Ahelegbey et al. (2012) the authors propose a Bayesian inferential approach, to analyze the dynamic interactions among macroeconomics variables in a graphical vector autoregressive model, that can be employed to overcome also the second problem.

The methodological contribution of this paper is to consider both the above extensions in a statistically correct estimation of the adjacency matrix of a financial network of countries. At first we will introduce a Bayesian (static) graphical model, able to overcome the first problem. Later we consider, as an extension of the first model, a Bayesian (dynamic) model, able to overcome both the first and the second problem. The results obtained from the application of different graphical models: classical, static Bayesian and dynamic Bayesian, will be compared in terms of adjacency and connectedness implications on a publicly available database, that contains cross-border financial flows. This will also allow us to compare conclusions in terms of central countries, most contagious in terms of systemic risk. We will also compare our conclusions with those resulting from the application of the Granger causality tests proposed in Billio et al. (2012), that can be seen as a special type of dynamic graphical model, therefore obtainable as a by-product of our approach.

The applied contribution of this paper is in the understanding of systemic risk. Systemic risk can be defined as the risk that the failure of one significant financial institution can cause or significantly contribute to the failure of other significant financial institutions as a result of their linkages to each other (see e.g. Billio et al., 2012). Researchers have recently proposed that network models can help model the systemic risk in financial systems which display complex degrees of connectedness. In particular, Billio et al. (2012) propose several econometric measures of connectedness based on principal component analysis and Granger-causality networks. They find that hedge funds, banks and insurance companies have become highly interrelated over the past decade, likely increasing the level of systemic risk through a complex and time-varying network of relationships. Chen et al. (2012) follows a similar line of research, using high frequency market value data on credit default swap spreads in the insurance sector. More recently, Barigozzi and Brownlees (2013) developed an algorithm based on a two step lasso regression that allows to estimate long

run sparse partial correlation networks. Their methodology is applied to a panel of U.S. bluechips, whose monthly equity returns are decomposed into a systematic and an idiosyncratic components. They found that the network structure of the idiosyncratic part is found to be statistically significant in explaining the variation of returns. In the previous papers, the transmission of contagion among financial agents depends on the degree of interconnectivity. This will also be true in our paper. In the previous papers, financial agents are private financial institutions: banks, hedge funds, insurance companies, listed companies in general. Differently, in our paper, we estimate the structure of contagion between country specific financial systems. The lack of existing information on the degree of connectedness among different countries remains a concern for regulatory bodies such as the Financial Stability board, the Financial Stability Oversight Council and the International Monetary Fund. Our contribution aims to provide a contribution in this direction.

The paper is organized as follows. Section 2 introduces graphical network models based on partial correlation matrices, and compares them with network models based on marginal correlations. In section 3, we introduce Bayesian graphical models, and show their theoretical implications: in particular, we describe a model averaging context and show how a dynamic approach can be built by decomposing the model into multivariate autoregressive and contemporaneous networks. In the same section, for the sake of completeness, we also introduce a network model based on Granger-causality tests, as the one proposed in Billio et al. (2012). Section 4 describe the empirical results obtained with the application of the previous network models to a set of Bank of International Settlement data that emphasize the role that countries have in terms of cross-border interlikages. Finally, section 5 contains some concluding remarks and future research directions.

2 Graphical models networks

Correlation based networks are suitable tools to visualize the structure of pairwise cross correlations among a set of N time series. If we associate different time series with different nodes of a network, each pair of nodes can be thought to be connected by an edge, with a weight that can be related to the correlation coefficient between the two corresponding time series. Thus, a network of N nodes can be described by its associated matrix of weights, named adjacency matrix, an $N \times N$ matrix, say A , with elements $a_{i,j}$. Alternatively, if the aim of the research is to focus on the structure of the interconnections, and less on their magnitude, the adjacency matrix can be made binary, setting $a_{i,j} = 1$ when two nodes are correlated and $a_{i,j} = 0$ when they are not correlated.

In both cases, network statistics can be derived using A : in particular, central-

ity measures can be obtained using an appropriate singular value decomposition of such matrix. The centrality measure that has been proposed in financial network modeling, to explain the capacity of an agent to cause systemic risk, that is, a large contagion loss on other agents, is the eigenvector centrality (see e.g. Furfine, 2003 and Billio et al., 2012). The eigenvector centrality measure is a measure of the importance of a node in a network. It assigns relative scores to all nodes in the network, based on the principle that connections to few high scoring nodes contribute more to the score of the node in question than equal connections to low scoring nodes.

More formally, for the i -th node, the centrality score is proportional to the sum of the scores of all nodes which are connected to it, as in the following equation:

$$x_i = \frac{1}{\lambda} \sum_{j=1}^N a_{i,j} x_j,$$

where x_j is the score of a node j , $a_{i,j}$ is the (i, j) element of the adjacency matrix of the network, λ is a constant and N is the number of nodes of the network.

The previous equation can be rewritten for all nodes, more compactly, as:

$$Ax = \lambda x,$$

where A is the adjacency matrix, λ is the eigenvalue of the matrix A , with associated eigenvector x , an N -vector of scores (one for each node).

Note that, in general, there will be many different eigenvalues λ for which a solution to the previous equation exists. However, the additional requirement that all the elements of the eigenvector be positive (a natural request in our context) implies (by the Perron–Frobenius theorem) that only the eigenvector corresponding to the largest eigenvalue provides the desired centrality measures. Therefore, once an estimate of A is provided, network centrality scores can be obtained from the previous equation, as elements of the eigenvector associated to the largest eigenvalue.

Centrality measures have been the object of growing attention, ever since network regimes were identified as the underlying structures of important phenomena found in many different fields. These measures have been typically applied to fully connected networks, based on adjacency matrices that do not contain zeros. In real situations, however, the adjacency matrix is typically sparse (it contains zeros), rather than fully connected. This implies the need to develop network models different from correlation based ones. This task has been accomplished through the construction of a Minimal Spanning Tree, as in Araújo and Mendes (2000) and Spelta and Araújo (2012).

Here we propose a method that is more widely applicable, being still based

on easy to understand correlation measures, but of a different kind. Our proposal is based on the following intuition. Correlations take into account the total effect of an explanatory variable on a dependent variable, without distinguishing between direct and indirect effects. To capture direct effects, partial correlations should be employed. A partial correlation coefficient measures the correlation between any two variables, when the remaining variables are held constant. In so doing only the direct correlation between two agents is measured, rather than the indirect one, that may arise from a common dependence on other variables.

From a statistical viewpoint, while correlations can be estimated, on the basis of the N observed time series of data, assuming that, at each time point, observations follow a multivariate Gaussian model, with unknown variance-covariance matrix Σ , partial correlations can be estimated assuming that the same observations follow a graphical Gaussian model, in which Σ is constrained by the conditional independence described by a graph (see e.g. Lauritzen, 1996, and Whittaker, 1990).

Let $X = (X_1, \dots, X_N) \in R^N$ be a random vector distributed according to a multivariate normal distribution $\mathcal{N}_N(\mu, \Sigma)$. In this paper, without loss of generality, we will assume that the data are generated by a stationary process, and, therefore, $\mu = 0$. In addition, we will assume throughout that the covariance matrix Σ is non singular.

Let $G = (V, E)$ be an undirected graph, with vertex set $V = \{1, \dots, N\}$, and edge set $E = V \times V$, a binary matrix, with elements e_{ij} , that describes whether pairs of vertices are (symmetrically) linked between each other ($e_{ij} = 1$), or not ($e_{ij} = 0$). If the vertices V of this graph are put in correspondence with the random variables X_1, \dots, X_N , the edge set E induces conditional independence on X via the so-called Markov properties (see e.g. Lauritzen (1996)).

More precisely, the pairwise Markov property determined by G states that, for all $1 \leq i < j \leq N$,

$$e_{ij} = 0 \iff X_i \perp X_j | X_{V \setminus \{i,j\}};$$

that is, the absence of an edge between vertices i and j is equivalent to independence between the random variables X_i and X_j , conditionally on all other variables $X_{V \setminus \{i,j\}}$.

In our context, all random variables are continuous and it is assumed that $X \sim \mathcal{N}_N(0, \Sigma)$. Let the elements of Σ^{-1} , the inverse of the variance-covariance matrix, be indicated as $\{\sigma^{ij}\}$. It can be shown that the following equivalence also holds:

$$X_i \perp X_j | X_{V \setminus \{i,j\}} \iff \rho_{ijV} = 0$$

where

$$\rho_{ijV} = \frac{-\sigma^{ij}}{\sqrt{\sigma^{ii}\sigma^{jj}}}$$

denotes the ij -th partial correlation, that is, the correlation between X_i and X_j conditionally on the remaining variables $X_{V \setminus \{i,j\}}$.

Therefore, by means of the pairwise Markov property, given an undirected graph $G = (V, E)$, a graphical Gaussian model can be defined as the family of all N -variate normal distributions $\mathcal{N}_N(0, \Sigma)$ that satisfy the constraints induced by the graph on the partial correlations, as follows:

$$e_{ij} = 0 \iff \rho_{ijV} = 0$$

for all $1 \leq i < j \leq N$.

In practice, the available data will be used to test which partial correlations are significantly different from zero, once a significance level threshold α is chosen. For the non significant partial correlations, the corresponding elements of the edge set E will be set to zero, thereby obtaining an estimate of the graphical structure G (being V fixed). Note that, in the context of network modelling, if the adjacency matrix is assumed to be binary, $A = E$ and, therefore, the selection of G gives automatically the adjacency matrix and the network centrality measures.

This way of proceeding is fine when the selected model is strongly supported, both in terms of data likelihood and in terms of coherence with economical subject-matter assumptions. When this is not the case, a more general approach, that takes model uncertainty into account, is needed. This is the aim of the next section.

3 Bayesian Graphical models networks

3.1 Static models

Graphical model uncertainty can be taken into account, along with parameter uncertainty, within a Bayesian approach, whose main practical advantage is that inferences on quantities of interest can be averaged over different models, each of which has a weight that corresponds to its Bayesian posterior probability. See, for example, Madigan et al. (1994), Giudici and Green (1999) and Castelo and Giudici (2001).

To achieve this aim, the first task is to recall the expression of the marginal

likelihood of a graphical Gaussian model, and specify prior distributions over the parameter Σ as well as on the graphical structures G .

For a given graph G , let $\pi^i(G)$ be the neighbors of a variable X^i , $i = 1, \dots, N$: the set of all vertices that are connected to i or, equivalently, the set of all random variables on which X^i is conditionally dependent, with partial correlations significantly different from zero. Statistical inferences can be based on the likelihood of the observed time series. Assuming that the observed data can be arranged in a matrix X , with n rows that are independent and identically distributed multivariate time series of dimension N , each following a multivariate normal distribution $\mathcal{N}_N(0, W)$, (where $W = \Sigma^{-1}$ is the precision matrix), the likelihood of the unknown parameters W, G is equal to:

$$P(X | W, G) = \prod_{t=1}^n \prod_{i=1}^N P(X_t^i | \pi^i(G), W)$$

As expected, the likelihood depends on the parameters W and G . According to the conventional Bayesian model averaging paradigm (see e.g. Madigan and York, 1995), we assume that both W and G are random, and that a joint prior distribution over (W, G) can be expressed by a discrete probability distribution over all graph structures G and, conditionally on each possible graph, by a continuous distribution over the set of parameters W :

$$P(G, W) = P(G)P(W | G)$$

For $P(G)$ we assume a uniform prior over all possible graphical structures while for the prior on the parameters we assume a conjugate Wishart prior, as in Giudici and Green (1999).

More formally, based on the assumption that we have available a random sample of size n from a multivariate normal distribution $\mathcal{N}_N(0, W)$, we assume that $P(W)$ is a Wishart distribution with $\alpha > N$ degrees of freedom and a scale matrix T_0 .

According to the Bayesian prior-to-posterior paradigm (see Bernardo et al., 1996) it can be shown that, under the previous assumptions, the posterior distribution of the precision matrix W , $P(W|X)$ is a Wishart distribution with $\alpha + N$ degrees of freedom and a scale matrix given as follows:

$$T_n = T_0 + S_n$$

where S_n is the sample variance-covariance matrix.

Note that the proposed prior distributions can also be used to integrate the likelihood with respect to the unknown random parameters, obtaining the so-

called marginal likelihood of a graph, which will be the main metric for model selection. This follows from

$$P(X | G) = P(X | W, G) P(W)$$

It can be shown that such marginal likelihood is equal to:

$$P(X | G) = \pi^{-\frac{N+n}{2}} \frac{c(N, \alpha + n)}{c(N, \alpha)} |T_0|^{\frac{\alpha}{2}} |T_n|^{-\frac{\alpha+n}{2}}$$

where $|T_0|$ and $|T_n|$ are the determinants of the matrices T_0 and T_n , respectively and $c(\cdot)$ is a normalization constant, given by:

$$c(x, y) = \prod_{j=1}^x \Gamma\left(\frac{y+1-j}{2}\right)$$

The metric expressed by the above marginal likelihood is the basic ingredient for graphical Gaussian model selection and averaging, as will now be shown.

According to the conventional Bayesian paradigm, being the model space discrete, the best graphical model will be that with the highest a posteriori probability.

By Bayes rule, the posterior probability of a graph is given by:

$$P(G | X) \propto P(X | G) P(G)$$

and, therefore, since we assumed a uniform prior over the graph structures, maximizing the posterior probability is equivalent to maximizing the marginal likelihood metric. For graphical model selection purposes we shall thus search in the space of all possible graphs for the structure such that

$$G^* = \arg \max_G P(G | X) \propto P(X | G).$$

The Bayesian paradigm does not force conditioning inferences on the (best) model chosen. The assumption of G being random, with a prior distribution on it, allows any inference on quantitative parameters to be model averaged with respect to all possible graphical structures, with weights that correspond to the posterior probabilities of each graph. This because, by Bayes' theorem:

$$P(W|X) = P(W|X, G)P(G|X)$$

The above allows to overcome the main drawback of non-Bayesian graphical models, namely, the fact that, once a model is chosen, all inferences will

be conditional on that model, even if it has a little support from the data (although maximal).

However, in real situations, the number of possible graphical structures may be very large and we may need to restrict the number of models to be averaged. This can be done efficiently, for example, following a simulation-based procedure for model search, such as Markov Chain Monte Carlo (MCMC) sampling, described in Madigan and York (1995). One of the standard MCMC methods is the Metropolis-Hastings (MH) algorithm, which is based on an acceptance-rejection scheme. In our context, given an initial graph, the algorithm samples a new graph using a proposal distribution. The newly sampled graph is then compared with the old graph, with a decision rule to either reject or accept the proposed sample. More precisely, following Giudici and Green (1999), the algorithm randomly selects a node from the current graph (G_{old}) and proposes an action to either add or delete a single edge to produce a new graph (G_{new}). The proposed graph G_{new} is either accepted, becoming (G_{old}) in the next step, or rejected in which case the previous graph is maintained as G_{old} . The decision to accept or reject a proposed graph depends on an acceptance probability.

By assuming a uniform graph prior, it can be shown that the log acceptance probability is given by:

$$\log(Ac) = \max \{ \log(P(X | G_{new})) - \log(P(X | G_{old})), 0 \}$$

Once the acceptance probability is calculated, we draw a random number U from a uniform distribution on $(0, 1)$. Then $\log(U)$ is compared with log acceptance probability. If $\log(U) < \log(Ac)$ the new proposal is accepted and added to the chain, otherwise the current structure is retained. Thus the mechanism automatically accepts samples showing improvements (i.e when $\log(Ac) > 0$) and accepts the rest with the acceptance probability (Ac).

3.2 *Dynamic models*

The Bayesian graphical model presented so far is a static model, that assumes that each of the N time series is made up of n independent and identically distributed observations. We now extend the approach in a more realistic dynamic setting. Following the idea of Ahelegbey et al. (2012), we build a graphical model that is made up of two parts: a simultaneous cross-sectional component, similar to the model in the previous subsection, and a novel multivariate dynamic autoregressive component.

To achieve this aim we recall the definition of a vector autoregressive (VAR) process. A VAR process of order s is of the form:

$$X_t = B_0 X_t + B_1 X_{t-1} + \dots + B_s X_{t-s} + \varepsilon_t \quad t = 1, \dots, n$$

where X_t is an N dimensional vector of time series realizations at time t , ε_t is an N dimensional vector of independent and serially uncorrelated structural disturbances with mean zero and a diagonal matrix Σ , and B_0, \dots, B_s are $N \times N$ regression matrices.

A dynamic graphical model (DGM) can be built exploiting the above vector autoregressive (VAR) representation of multivariate time series observations. In a VAR model, for any given time lag s , we can establish a one-to-one correspondence with a graphical model, by setting a zero edge between two vertices i and j whenever, for any given time lag s , a dependent variable X^j , observed at time t , is independent from an explanatory variable X^i , observed at time $t - s$ as follows:

$$e_s^{i-j} = 0 \implies X_t^i \perp X_{t-s}^j \mid \left(X_t^{V \setminus \{i\}}, X_{t-s}^{V \setminus \{j\}} \right)$$

We can exploit equivalences as the above one to build a dynamic graphical model, for a specified time lag. More formally, a DGM is a pair $(G, G \downarrow)$, where G is a graphical model, that defines the contemporaneous conditional dependences between the available random variables; and $G \downarrow$ is another graphical model, that defines the temporal conditional dependences between the same variables, at different times. In $G \downarrow$, the neighbors of a node are in previous time periods, and are conventionally named “the parents”, because of their implied asymmetric relationship with the “child” node, that is supposed to be caused by them.

A VAR process of order s and, correspondingly, a dynamic graphical model of order s , assume that the within period (contemporaneous) conditional dependence among variables are described by a Multivariate instantaneous network (MIR) graphical component. In addition, the lag s conditional dependency structure between the variables is constant in time, depends only on the lag s and is described by a Multivariate AutoRegressive network (MAR) component. Assume, from now on, for simplicity and without loss of generality, that $s = 1$.

Assume that we have now available, for each time point $t = 1, \dots, n$, a random observation from a multivariate normal distribution $\mathcal{N}_N(0, W)$, (where $W = \Sigma^{-1}$ is the precision matrix).

Let π_{t-1}^i and π_t^i be the neighbors of X_t^i in the MAR and MIR networks. Following Ahelegbey et al. (2012), it can be shown that the marginal likelihood function decomposes according to the structure of the model into a MIR and a MAR component, as follows:

$$\begin{aligned}
P(X | W, G) &= \prod_{t=1}^n \prod_{i=1}^N P\left(X_t^i | \pi_{t-1}^i(G \downarrow), \pi_t^i(G), W\right) = \\
&= \prod_{t=1}^n \prod_{i=1}^N P\left(X_t^i | \pi_{t-1}^i(G \downarrow), W\right) \times \prod_{t=1}^n \prod_{i=1}^N P\left(X_t^i | \pi_t^i(G), W\right)
\end{aligned}$$

This decomposition of the structure facilitates the inference procedure such that we can learn the MIR network independently from the MAR network. Model search simplifies into searching for the network that maximizes each marginal likelihood score independently, according to what shown in the previous subsection.

In this paper, given the high dimensionality of the model space we consider MCMC sampling approximate computations. The design of our Monte Carlo sampling for the MAR network takes inspiration from what proposed in Grzegorzczak (2010). Specifically, the observed multivariate time series are transformed into $1 \times n$ data cells with each cell composed of $(N + 1)(n - 1)$ time series matrix. Each cell corresponds to a transformed data by extracting the series of the i -th variable and shifting it one time ahead of the other variables. That is the last observation of the original data is deleted to obtain a matrix of $N \times (n - 1)$. The posterior scale matrix T_n for MAR is therefore computed for each cell by replacing N with $N + 1$ and n with $n - 1$.

3.3 Granger causality models

We conclude this section with a remark on a special type of dynamic graphical model, that coincides with the Granger-causality procedure proposed in Billio et al. (2012). Granger-causality tests can be used to identify a network of statistically significant Granger-causal relations among economical agents. They are aimed at establishing whether a relationship between past values of one variable and future values of another one does or does not hold.

In addition, Granger-causality test help investigating the directionality of such relationships: a time series j is said to “Granger-cause” time series i if past values of j contain information that helps to predict i .

From a mathematical view point, the formulation of Granger-causality tests is based on linear regressions of R_{t+1}^i on R_t^i and R_t^j . The series are supposed stationary with zero mean and the model of linear interrelationship can be formulated, following Billio et al. (2012) as:

$$\begin{aligned}
R_{t+1}^i &= a^i R_t^i + b^{ij} R_t^j + e_{t+1}^i \\
R_{t+1}^j &= a^j R_t^j + b^{ji} R_t^i + e_{t+1}^j
\end{aligned}$$

with e_{t+1}^i and e_{t+1}^j uncorrelated zero mean gaussian random variables.

Then, j Granger-causes i when b^{ij} is different from zero. On the other hand i Granger-causes j when b^{ji} is different from zero. When both of these statements are true, there is a feedback relationship between the time series. All of these alternative statements can be verified within a dynamic graphical modeling context, evaluating whether the three possible edges between i and j are different from zero.

Thus, Granger causality networks are a special type of graphical dynamic models. However, Granger testing is usually performed in a non-Bayesian context and, therefore, the corresponding results are less robust than those based on the Bayesian graphical dynamic models presented in the last subsection.

4 Empirical analysis

4.1 Data

The Bank for International Settlements locational banking statistics (LBS), that includes aggregate international claims and liabilities of reporting banks by country of residence, provides a plentiful data set of aggregate cross-border exposures for a set of reporting and non-reporting countries all over the world.

The main purpose of the LBS statistics is to provide information on the role of banks and financial centers in the intermediation of international capital flows. They were originally intended to complement monetary and credit aggregates, being consistent with both the national balances of payments and the systems of national accounts. LBS statistics are based on quarterly data on the gross international financial claims and liabilities of banks resident in a given country, excluding money market funds. LBS statistics were originally introduced in 1964 to monitor the development of euro-currency markets, although in practice comprehensive locational data are available since the 1980s. Here we consider 24 countries, reported in Table 1, for which the data are complete and reliable. Each country is represented by the value of its liabilities *vis-a-vis* the other reporting countries, measured on a quarterly basis, from the last quarter of 1983 (Q4-1983) to the third quarter of 2011 (Q3-2011).

AT: Austria	IT: Italy
BS: Bahamas	JP: Japan
BH: Bahrain	LU: Luxemburg
BE: Belgium	NL: Netherlands
CA: Canada	AN: Netherlands Antilles
KY: Cayman Islands	NO: Norway
DK: Denmark	SG: Singapore
FI: Finland	ES: Spain
FR: France	SE: Sweden
DE: Germany	CH: Switzerland
HK: Hong Kong	GB: United Kingdom
IE: Ireland	US: United States

Table 1: Reporting Countries

4.2 Marginal correlation networks

We first consider the application of classical network models, based on marginal correlations, to LBS data. The 24×24 adjacency matrix (A), with elements $a_{i,j}$, can be obtained on the basis of a 24×24 correlation matrix, for the set of 24 reporting countries ($N = 24$), calculated on $n = 110$ time observations. Instead of using a fully connected network, as in the network modeling literature (see e.g. Araújo and Mendes, 2000), we consider a “statistical” network, in which the edge that connects two countries is present on the basis of a pairwise F-test, that informs whether the corresponding correlation is significant or not, with a significance level α .

Figure 1 shows the network obtained on the basis of such an adjacency matrix, taking a significance level equal to $\alpha = .10$.

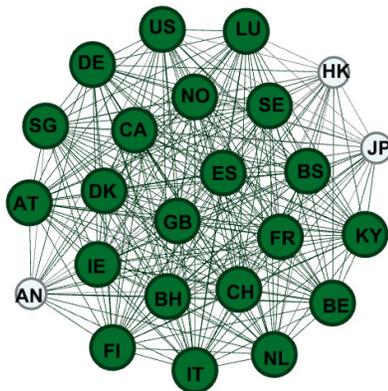


Figure 1: marginal correlations network

From Figure 1, note that the network is not fully connected, even if only few links are removed. More precisely, most of the countries have all the 23 links with the others while JP and HK 22 and AN only 21. These latter countries seem, therefore, relatively less correlated with the rest of the financial system being considered.

Economically, our results above may be interpreted attributing to JP, HK and AN (but especially to the latter two) a role of “countercyclical buffers”, less subject to financial cycles. Indeed, according to Errico and Borrero (1999), the role that the Offshore Financial Centers (OFC), such as HK and AN, play in the allocation of financial capital worldwide has grown, for example during the Asian and Argentina crises in the earlier nineties. A clarifying contribution comes from the paper of Huizinga and Nicodeme (2004), who examine the relation between tax policies and the amount of foreign liabilities in each national banking sector. Foreign banks are expected to have relatively abundant opportunities to shift profits when the host countries are characterized by low levels of taxation. In this context, they found enough evidence on the inward profit shifting role played by Hong Kong and the Netherlands Antilles.

4.3 *Partial correlations network*

We now apply classical graphical Gaussian model to LBS data, and derive an adjacency matrix based on graphical model selection. In such adjacency matrix, two countries will be linked if the corresponding edge is present in

the selected graphical model or, equivalently, when the corresponding partial correlation is equal to zero.

Figure 2 shows the network obtained on the basis of the selected graphical model, taking a significance level equal to $\alpha = .10$. The size and the color of the nodes is a function of the number of edges that point on a node. Dark and thick nodes are the ones with the highest degree, that is, most linked with respect to the other variables.

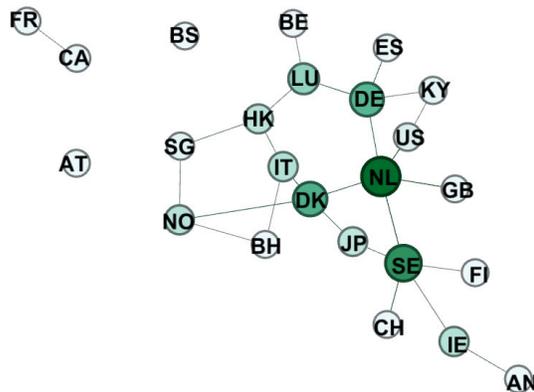


Figure 2: partial correlation based network

From Figure 2 we can see that the selected graph and, therefore, the corresponding adjacency matrix, is rather sparse, especially in comparison with the network in Figure 1. Indeed, differently from before, the average number of edges pointing to a node is 2.083: each node is connected, on average, to only two other nodes. This can be explained recalling the difference between marginal and partial correlations: while marginal correlations are unconditional and reflect all comovements between two variables (direct and indirect), partial correlations are conditional on the dependences described by the selected graph and measure only direct correlations. So, while marginal correlations provide an adjacency matrix that is too connected, partial correlations provide a too sparse one. A representation that is “intermediate” between these two will be provided by Bayesian graphical models, in the next subsections.

Before moving to the application of Bayesian graphical models, we can nevertheless draw some preliminary economical conclusions from Figure 2. For example, note that country with the highest number of connecting edges is NL, followed by SE, DK and DE. It is well known that the Dutch financial system is largely exposed to the rest of the world, also in a direct way, having large banks that operates at a high cross-border level. On the other hand, the

role of DK, SE could be explained by the crisis that many northern countries suffered together at the beginning of the nineties, and that of DE by the recent European financial crisis. Indeed, the first banking system touched by the subprime crisis was the German banking system, in which IKB was the first European bank affected.

A second remark is that the US, as stressed by Von Peter (2007), has few connecting edges, as its correlation is with all the system, and few countries in particular. Similarly, GB, that is another important financial hub, is also little directly connected (only with NL).

4.4 *Static Bayesian networks*

We now present the results from the application of the Bayesian graphical model. The main advantage of the Bayesian approach to graphical models is the possibility of model averaging the results obtained with single models, with weights provided by the corresponding graph posterior probabilities. This idea can be applied to the adjacency matrix elements, which, therefore, become relative frequencies of edge presence. Figure 3 presents the network obtained with such an adjacency matrix. In order to obtain Figure 3 we have run our MCMC algorithm for 8500 iterations, using the last 250 iterations to calculate edge presence frequencies.

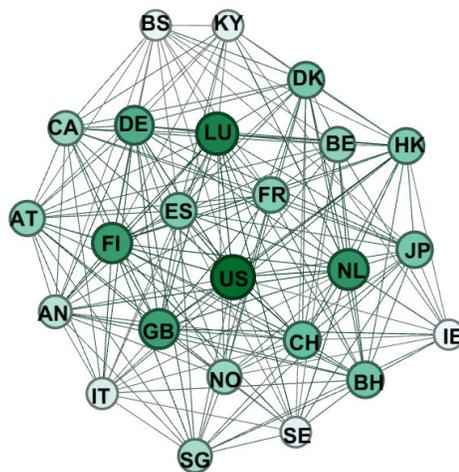


Figure 3: static Bayesian graphical network

From Figure 3, the Bayesian graphical model suggests that LU and US, followed by GB, NL, FI and DE are central nodes. To better understand their

relative importance, we can calculate the eigenvector centrality measure. It turns out that LU and US display the highest centrality measure, followed by FI, DE, NL, GB, followed by CH.

While the role of the United States and that of the United Kingdom can be explained by the role these two countries have in the world financial network, as international hubs (see Minoiu and Reyes, 2011), the position achieved by the Netherlands, Switzerland, and also that of Germany can be attributed to their large cross-border exposures which reflects into their high procyclicality with the rest of the international system. The two situations are indeed distinct: while the position of the United States and that of the United Kingdom reflects their position as a host to many foreign banks, countries like the Netherlands, Germany and Switzerland are home to multinational banks generating considerable interoffice activity across borders (see Von Peter, 2007).

Finally, the position of Luxembourg and that of Finland may be attributed to a global or local role as “off-shore” countries. Luxembourg brings in less than 44 billion euros from its goods and services, yet its banks boast 227 billion euros in deposits. The international financial crisis nearly threw into bankruptcy the two biggest private banks of the country, Dexia-Bil, the oldest bank of the Grand Duchy, and Fortis, formerly Banque Générale du Luxembourg. Finland is not a traditional international financial center, however, banks in Finland are permitted to open account for non-resident foreigners. In the early 1980s the financial market was mostly deregulated, leading to a massive credit expansion largely based on foreign debt. This turmoil culminated with the collapsed on September 1991 of the Skopbank, and the took over of the Bank of Finland.

To complete the report of the static Bayesian analysis, Figure 4 contains three different diagnostics of convergence of MCMC simulations, based, respectively, on the number of edges present in the estimated model (Figure 4a) on the estimated log-likelihood (Figure 4b) and on the cumulative difference between the number of accepted and rejected models (Figure 4c).

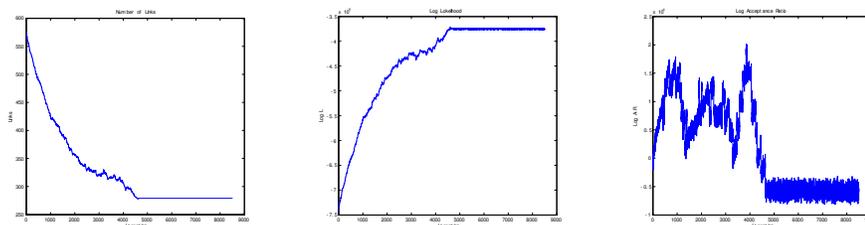


Figure 4: convergence diagnostics

In Figure 4, and in each of the three subplots, the x -axis represents the running iteration. Note that, after about the first 4500 iterations, the Markov

chain starts to converge, according to all three diagnostics.

4.5 Dynamic Bayesian networks

We now consider the application of the dynamic Bayesian model. Figure 5 shows the adjacency matrix based on the edge presence frequencies estimated by model averaging the binary edge indicators over all selected models. Figure 6 shows the resulting centrality measures. For the sake of comparison, we have used the same MCMC settings as before. Of course, the number of nodes is now higher ($24 + 24 = 48$) and, therefore, so is the number of possible edges that connect each node: $23 + 23 = 46$ for the “contemporaneous” nodes, as they can be affected by any of the other countries at the same time as well as in the previous time period. In contrast, the 24 lagged nodes have at most 23 connections as, by definition, a node in the MAR network can be linked to a mode in a subsequent time but not to a node at the same time. In figure 5, therefore, the lagged nodes will appear “lighter” than the others.

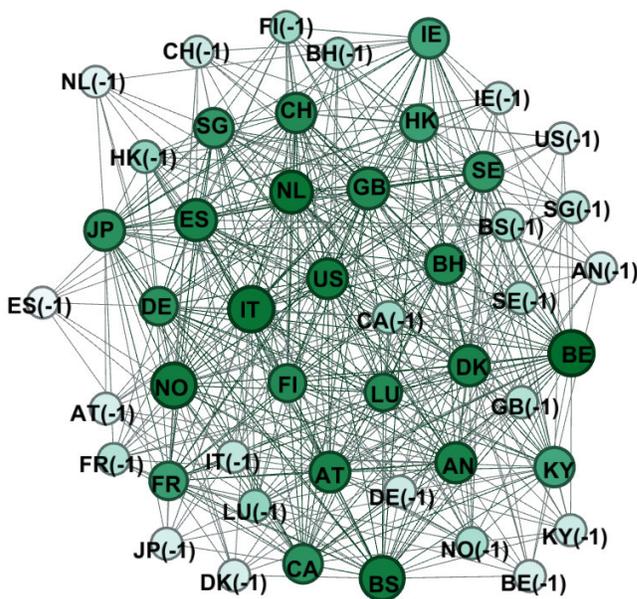


Figure 5: dynamic Bayesian graphical network

On the basis of Figure 5, and the associated centrality measures displayed by Figure 6, the countries that are most central in the MIR component of the dynamic model are the US, followed by NL, LU and FI as in the Bayesian Static model.

If we focus on lagged variables in the MAR component, the highest centrality is shown by LU, HK and FI even if the score is lower, as we expect by definition,

when compared with the contemporaneous variables. The role of off-shores countries is therefore recovered through the lagged variables.

Comparing the overall dynamic model with respect to static one, note the presence of BE in the countries that own the highest centrality. This evidence can be explained following Garratt et al. (2011), who claims that Belgium and the Netherlands have become heavily interdependent. Furthermore there appear a lower centrality for DE, CH and the UK which are overtaken by two southern European countries IT, ES.

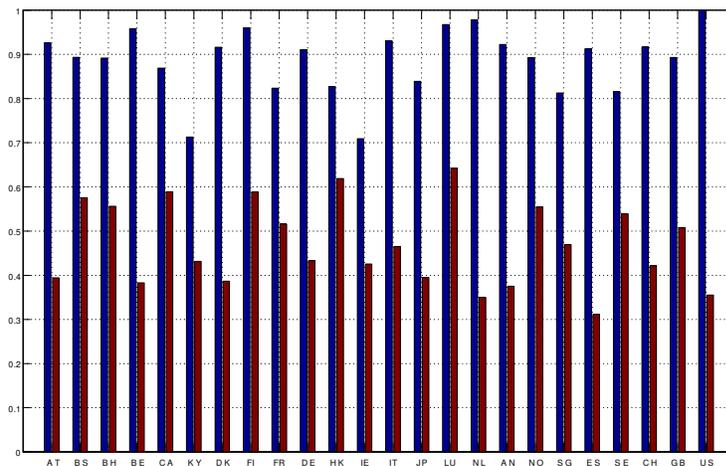


Figure 6: eigenvector centrality (blue-MIR)-(red-MAR)

As before we complete the analysis with the diagnostics of convergence. Figure 7 shows the results.

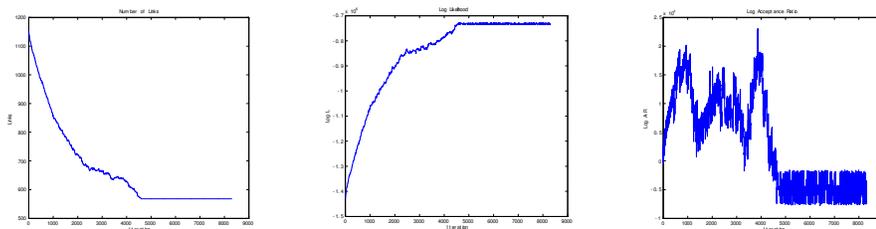


Figure 7: convergence diagnostics

Figure 7 shows the convergence of the chosen diagnostics around 4550 iterations, for all of the three plotted convergence measures.

4.6 Granger causality networks

For the sake of comparison we now consider a Granger-causality network, built as described in the theoretical section. Figure 8 shows the application of such model to the LBS statistics.

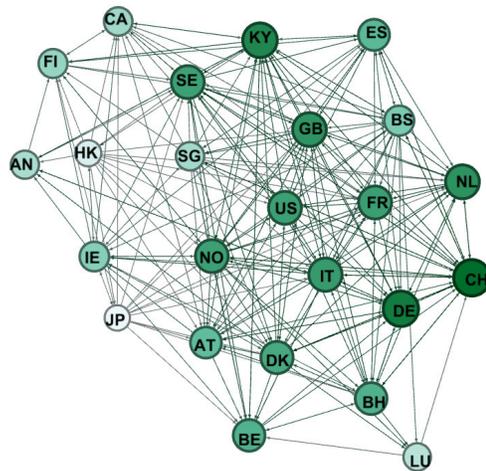


Figure 8: Granger causality based network

The network in Figure 8, and the associated centrality measures, emphasize the role of the European countries, that are, on average, more central than non European ones. In particular, the Asian countries are, on average, the most peripheral. The biggest hubs, according with the eigenvector centrality measure, are CH, DE and NL: this can be explained, as in the static model, by their multinational banks that generate considerable interoffice activity across borders (see also Von Peter, 2007). They are followed by KY and than by NL, GB, the US and IT similarly as in the Bayesian dynamic model. The novelty is the role of the Cayman Islands (KY), clearly an off-shore country as are LU, HK and FI in the dynamic model. Indeed as pointed out by Von Peter (2007), the Cayman Island concentrate most of their positions on US banks, which is therefore a related node. In addition, we find also that GB can explain the movement in KY banks liabilities.

We remark that the Granger causality network is a special case of the dynamic models that we have considered previously. The main difference is that, while results in our dynamic models are model averaged, the results from Granger testing are conditional on a single model and, therefore, less robust and reliable than those based on dynamic Bayesian networks. However, the comparison of the results indicates similar main economical interpretations, in terms of countries that are more central and, those, more problematic in a systemic risk framework.

5 Conclusions

Network models are a useful tool to model interconnectedness and systemic risks in financial systems. Such models are essentially descriptive, and based on highly correlated networks.

The motivation of this paper is to provide a stochastic framework for financial network models, aimed at selecting only the statistically significant relationships, leading to a parsimonious and sparse representation of the network.

The paper contains two main research contributions in this direction. First we introduce multivariate Gaussian graphical models, defined in terms of Markov properties, on the basis of which estimation of the adjacency matrix can be performed statistically testing conditional independences, equivalent to specified zeroes among certain (partial) correlation coefficients. Second, we have proposed to robustify graphical model selection by means of a Bayesian approach, both in a static and in a time-varying framework, thus allowing model uncertainty to be taken into account in the estimation of the adjacency matrix. The dynamic version provides a convenient framework to model temporal data.

We have applied our proposed methods to the Bank of International Settlements locational banking statistics, with the aim of identifying central countries, whose financial systems could have a high degree of interconnectedness, whose failures could result in further distress or breakdowns in the whole system.

Our results show that the countries that are potentially most contagious can be splitted in three main groups: international financial hubs such as US and GB; off-shore countries such as LU, HK, FI and KY and, finally, countries with large cross-border financial activities as NL, CH and DE. A fourth group of countries, including weak financial systems, emerges only when dynamic lagged effects are properly considered, as with a dynamic Bayesian network model.

Future research include the extension of what proposed to non Gaussian graphical models and the application of what proposed here to the study of the interconnectedness between institutions, rather than countries: banks, insurance companies and hedge funds, as in Billio et al. (2012).

More generally, the field of financial systemic risk modelling is an area where good statistical thinking and statistical analysis can contribute substantially to reach conclusions that are more and more important and timely for policy makers, both at the national and international levels.

References

- [1] Ahelegbey, D. F., Billio, M., & Casarin, R. (2012). Bayesian Graphical Models for Structural Vector Autoregressive Processes. University Ca'Foscari of Venice, Dept. of Economics Research Paper Series No, 36.
- [2] Araújo, T., & Mendes, R. V. (2000). Function and form in networks of interacting agents. arXiv preprint nlin/0009018.
- [3] Barigozzi, M., & Brownlees, C. (2013). Nets: Network Estimation for Time Series. Available at SSRN.
- [4] Bernardo, J. M., Berger, J. O., Dawid A. P. & A. F. M. Smith (eds). (1996). Bayesian Statistics 5. Oxford: Oxford University Press.
- [5] Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104(3), 535-559.
- [6] Boss, M., Elsinger, H., Summer, M., & Thurner, S. (2004). An empirical analysis of the network structure of the Austrian interbank market. *Oesterreichische Nationalbank's Financial stability Report*, 7, 77-87.
- [7] Castrén, O., & Kavonius, I. K. (2009). Balance sheet interlinkages and macro-financial risk analysis in the euro area.
- [8] Chen, H., Cummins, J. D., Viswanathan, K. S., & Weiss, M. A. (2013). Systemic risk and the interconnectedness between banks and insurers: An econometric analysis. *Journal of Risk and Insurance*.
- [9] Cont, R., Moussa, A., & Santos, E. (2011). Network structure and systemic risk in banking systems. Edson Bastos e, *Network Structure and Systemic Risk in Banking Systems* (December 1, 2010).
- [10] De Masi, G., Iori, G., & Caldarelli, G. (2006). Fitness model for the Italian interbank money market. *Physical Review E*, 74(6), 066112.
- [11] Edwards, D. (1990). Hierarchical interaction models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 3-20.
- [12] Errico, L., & Borrero, A. M. (1999). Offshore Banking-An Analysis of Micro-and Macro-Prudential Issues. *International Monetary Fund*.
- [13] Fujiwara, Y., Aoyama, H., Ikeda, Y., Iyetomi, H., & Souma, W. (2009). Structure and temporal change of the credit network between banks and large firms in Japan. *Economics: The Open-Access, Open-Assessment E-Journal*, 3, 7.
- [14] Furfine, C. H. (2003). Interbank exposures: Quantifying the risk of contagion. *Journal of Money, Credit and Banking*, 111-128.
- [15] Gai, P., & Kapadia, S. (2009). A network model of super-systemic crises. *Documentos de Trabajo (Banco Central de Chile)*, (542), 1.

- [16] Garratt, R., Mahadeva, L., & Svirydzenka, K. (2011). Mapping systemic risk in the international banking network.
- [17] Georg, C. P., & Poschmann, J. (2010). Systemic risk in a network model of interbank markets with central bank activity (No. 2010, 033). Jena economic research papers.
- [18] Grzegorzczuk, M. (2010). An introduction to Gaussian Bayesian networks. In *Systems Biology in Drug Discovery and Development* (pp. 121-147). Humana Press.
- [19] Giudici, P., & Green, A. P. (1999). Decomposable graphical Gaussian model determination. *Biometrika*, 86(4), 785-801.
- [20] Giudici, P. and Castelo, R. (2001). Association models for web mining. *Data mining and knowledge discovery*, 5(3), 183-196.
- [21] Helpman, E., & Krugman, P. R. (1985). *Market structure and foreign trade: Increasing returns, imperfect competition and the international economy*. The MIT press.
- [22] Huizinga, H., & Nicodème, G. (2004). Are international deposits tax-driven. *Journal of Public Economics*, 88(6), 1093-1118.
- [23] Jackson, M. O., & Watts, A. (2002). The evolution of social and economic networks. *Journal of Economic Theory*, 106(2), 265-295.
- [24] Kubelec, C., & Sá, F. (2010). The geographical composition of national external balance sheets: 1980-2005 (No. 384). Bank of England.
- [25] Lauritzen, S. L., & Wermuth, N. (1989). Graphical models for associations between variables, some of which are qualitative and some quantitative. *The Annals of Statistics*, 31-57.
- [26] Lauritzen, S. L. (1996). *Graphical models*. Oxford University Press.
- [27] Markose, S., Giansante, S., Gatkowski, M., & Shaghghi, A. R. (2009). Too interconnected to fail: Financial contagion and systemic risk in network model of cds and other credit enhancement obligations of us banks. analysis.
- [28] Madigan, D., York, J., & Allard, D. (1995). Bayesian graphical models for discrete data. *International Statistical Review/Revue Internationale de Statistique*, 215-232.
- [29] McGuire, P., & Tarashev, N. (2006). Tracking international bank flows. *BIS Quarterly Review*, 27-40.
- [30] Minoiu, C., & Reyes, J. A. (2013). A network analysis of global banking: 1978–2010. *Journal of Financial Stability*.
- [31] Nier, E., Yang, J., Yorulmazer, T., & Alentorn, A. (2007). Network models and financial stability. *Journal of Economic Dynamics and Control*, 31(6), 2033-2060.

- [32] Pearl, J. (1988). Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann.
- [33] Peltonen, T. A., Scheicher, M., & Vuillemeys, G., (2013). The Network Structure of the CDS Market and its Determinants. ECB, Working Paper Series NO 1583 / august 2013
- [34] Soramäki, K., Bech, M. L., Arnold, J., Glass, R. J., & Beyeler, W. E. (2007). The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications*, 379(1), 317-333.
- [35] Spelta, A., & Araújo, T. (2012). The topology of cross-border exposures: beyond the minimal spanning tree approach. *Physica A: Statistical Mechanics and its Applications*, 391(22), 5572-5583.
- [36] Von Peter, G. (2007). International banking centres: a network perspective. BIS Quarterly Review Working Paper.
- [37] Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *nature*, 393(6684), 440-442.
- [38] Wermuth, N., & Lauritzen, S. L. (1990). On substantive research hypotheses, conditional independence graphs and graphical chain models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 21-50.
- [39] Whittaker, J. (2009). *Graphical models in applied multivariate statistics*. Wiley Publishing.