Graphical network models for international financial flows

P. Giudici [1]

Department of Economics and Management, University of Pavia

A. Spelta [2]

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. To improve the statistical performance of financial network models, we propose to generate them by means of multivariate graphical models. We then 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.

[1] paolo.giudici@unipv.it

[2] alessandro.spelta01@universitadipavia.it

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

1 Introduction

The recent financial crisis has shown that financial markets and, more specifically, interbank markets, are highly interdependent. Network models can describe such interdependencies, as shown for example in Billio et al. (2012).

The network structure of national interbank markets has been studied by Soramaky and coauthors (2007), who analysed interbank payments transferred between commercial banks by the Fedwire Funds Services. A related contribution is the work of Fujiwara 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., 2006, Boss et al., 2004).

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 (2013). In particular, Minoiu and Reyes (2013) use confidential data representing cross-border bilateral financial flows intermediated by national banking systems, and found evidence of important structural changes in banking financial networks, following the occurrence of stress events. The same authors point out that their result should be interpreted with some caution because of the large amount of non-reporting countries (their sample contains 184 countries, of which only 15 report bilateral positions to the BIS).

Our aim is to extend Minoiu and Reyes (2013), using data on the total financial exposure of each country with respect to the rest of the world: a database that, besides being publicly available, is more reliable. Applying a stochastic network

model to such data, we establish bilateral links between countries, that can be used to understand which countries are most central and, therefore, most contagious (or subject to contagion).

To improve the performances of financial network models we introduce graphical models, that enrich the network perspective with a more structured statistical approach. As recently argued in Spelta and Araújo (2012), Billio et al (2012) and Barigozzi and Brownlees (2013), 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.

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 represent the variables and whose edges show 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).

From their appearance in the 90's, several methodological advances have been made for graphical models, but less so in terms of applications, especially in economics. This is because of two main problems, that require more advanced formulations.

First, 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 essentially 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 the second problem as well.

The methodological contribution of this paper is to consider both the above extensions in a financial network setting. From an applied viewpoint, the results obtained from the application of different graphical models will be compared in terms of adjacency matrix and implications on systemic risk transmission between countries.

The paper is organized as follows. Section 2 introduces financial networks based on graphical models, and compares them with non graphical ones. 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. Section 4 describes the empir-

ical results obtained with the application of the previous network models to the Bank of International Settlement cross-border financial flow data. Finally, section 5 contains some concluding remarks and future research directions.

2 Graphical models

Systemic risk can be represented by a network that describes the mutual relationships between the different economical agents involved.

Correlation based networks are suitable to visualize the structure of pairwise marginal 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.

It is well known that pairwise marginal correlations measure both the direct and the indirect effect of a variable on another. If the aim is to measure only the direct effect between two variables, without the "mediation" of others, pairwise partial correlations, rather than marginal ones, should be calculated.

From a statistical viewpoint, while correlations can be estimated, on the basis of N observed time series of data, assuming that 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 or, from an econometric viewpoint, Carvalho and West, 2007 and Corander and Villani, 2006).

Let $x = (x_1, ..., x_N) \in R^N$ be a $N-$dimensional random vector distributed according to a multivariate normal distribution $\mathcal{N}_N(\mu, \Sigma)$. 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 not singular.

Let $G = (V, E)$ be an undirected graph, with vertex set $V = \{1, ..., N\}$, 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 $X_1, ..., X_N$, 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 $1 \leq i < j \leq N$:

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

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 $\{\sigma^{ij}\}$.Whittaker (1996) proved that the following equivalence also holds:

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

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:

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

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

In practice, the available data will be used to test which partial correlations are different from zero, once a significance level threshold α is chosen. This leads to the selection of a graphical model on which to condition all inferences. In the next section we propose a Bayesian graphical model that takes model uncertainty into account.

Once a network is estimated, for example on the basis of a graphical Gaussian model, a natural request is to summarise it into a systemic risk measure. This request is quite reasonable, not only from a descriptive viewpoint, but also to provide an indicator that can act as an "early warning" predictive monitor.

The summary measure that has been proposed in financial network modeling to explain the capacity of an agent to cause systemic risk, that is, a large con-

tagion loss on other agents, is the eigenvector centrality (see e.g. Furfine, 2003 and Billio et al., 2012). The eigenvector centrality measures the importance of a node in a network by assigning 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 eigenvector centrality 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 eigenvectors 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.

Remark. The centrality measure is often applied to sparse networks, based on adjacency matrices that contain several zeros. This implies the need of establishing a criterion to decide whether the similarity between two agents is relevant enough to join them with a link. In the network literature, this task has been accomplished through the construction of a Minimal Spanning Tree (see for example Araújo and Mendes, 2000 and Spelta and Araújo, 2012). The application of the Minimal Spanning Tree criterion is, however, limited. With this method one is forced, a priori, to deal with a network that is a tree, therefore without cycles and with only $N-1$ edges. This construction therefore neglects part of the information contained in the adjacency matrix, since it only takes the $N-1$ edges that are considered in the hierarchical clustering process. The partial correlation networks that we have proposed, based on graphical Gaussian models, have instead a general applicability.

3 Bayesian Graphical models

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 (2003).

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, consider a sample X of size n from $P = \mathcal{N}_N(0, \Sigma)$, and let S_n be the corresponding observed variance-covariance matrix. For a subset of vertices $A \subset N$, let Σ_A denote the variance-covariance matrix of the variables in X_A, and define with S_A the corresponding observed variance-covariance submatrix.

To derive the likelihood of a graphical Gaussian model, we now recall the notion of a decomposable graph. Let G_A be a subgraph, obtained from the graph G considering as vertex set the elements of A and as edge set the subset of the edge set of G that contains only the vertices in A. The subgraph G_A is complete if all its vertices are joined by an edge. A complete subgraph that is not contained within another complete subgraph is called a clique. An ordering of the cliques of an undirected graph, $(C_1, ..., C_n)$ is said to be perfect if the vertices of each clique C_i, also contained in any previous clique $C_1, ..., C_{i-1}$, are all member of one previous clique; that is, for $i = 2, ..., n$,

$$S_i = C_i \cap \bigcup_{j=1}^{i-1} C_i \subseteq C_h$$

for some $h = h(i) \in \{1, 2, ..., i-1\}$. The sets S_i are called separators. If an undirected graph admits a perfect ordering it is said to be decomposable.

When the graph G is decomposable the likelihood of the data, under the graphical Gaussian model specified by P, nicely decomposes as follows (see e.g. Dawid and Lauritzen, 1993):

$$p(x|\Sigma, G) = \frac{\prod_{C \in \mathcal{C}} p(x_C|\Sigma_C)}{\prod_{S \in \mathcal{S}} p(x_S|\Sigma_S)},$$

where C and S respectively denote the set of cliques and separators of the graph G, and:

$$P(x_C|\Sigma_C) = (2\pi)^{-\frac{n*|C|}{2}}|\Sigma_C|^{-n/2}exp[-1/2tr\left(S_C\left(\Sigma_C\right)^{-1}\right)]$$

and similarly for $P(x_S|\Sigma_S)$.

Dawid and Lauritzen (1993) propose a convenient prior for the parameters of the above likelihood, the hyper inverse Wishart distribution. It can be obtained from a collection of clique specific marginal inverse Wisharts as follows:

$$l(\Sigma) = \frac{\prod_{C\in\mathcal{C}} l(\Sigma_C)}{\prod_{S\in\mathcal{S}} l(\Sigma_S)}$$

where $l(\Sigma_C)$ is the density of an inverse Wishart distribution, with hyperparameters T_C and α, and similarly for $l(\Sigma_S)$. For the definition of the hyperparameters here we follow Giudici and Green (1999) and let T_C and T_S be the submatrices of a larger matrix T_0 of dimension $N \times N$, and choose $\alpha > N$. To complete the prior specification, for $P(G)$ we assume a uniform prior over all possible graphical structures.

Dawid and Lauritzen (1993) show that, under the previous assumptions, the posterior distribution of the variance-covariance matrix Σ is a hyper Wishart distribution with $\alpha + N$ degrees of freedom and a scale matrix given by:

$$T_n = T_0 + S_n$$

where S_n is the sample variance-covariance matrix. This result can be used for quantitative learning on the unknown parameters, for a given graphical structure.

In addition, the proposed prior distribution can be used to integrate the likeli-

hood with respect to the unknown random parameters, obtaining the so-called marginal likelihood of a graph, which will be the main metric for structural learning, that involves choosing the most likely graphical structures.

Giudici and Green (1999) show that such marginal likelihood is equal to:

$$P\left(x|G\right)=\frac{\prod_{C\in\mathcal{C}}p(x_C)}{\prod_{S\in\mathcal{S}}p(x_S)}$$

where

$$p\left(x_C\right)=(2\pi)^{-\frac{n*|C|}{2}}\frac{k\left(|C|,\alpha+n\right)}{k\left(|C|,\alpha\right)}\frac{\det\left(T_0\right)^{\alpha/2}}{\det\left(T_n\right)^{(\alpha+n)/2}}$$

where $k(\cdot)$ is the multivariate gamma function, given by:

$$k_p\left(a\right)=\pi^{\frac{p(p-1)}{4}}\prod_{j=1}^{p}\Gamma\left(a+\frac{1-j}{2}\right).$$

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

$$P\left(G|x\right)\propto P\left(x|G\right)P\left(G\right)$$

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

$$G^*=\arg\max_G P\left(G|x\right)\propto\arg\max_G P\left(x|G\right).$$

The Bayesian approach 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 is due to Bayes' Theorem:

$$P(\Sigma|X) = P(\Sigma|x, G)P(G|x)$$

However, in many real problems, the number of possible graphical structures could 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).

In our context, given an initial graph, the algorithm samples a new graph using a proposal distribution. To guarantee irreducibility of the Markov chain, we follow Castelo and Giudici (2003) and test whether the proposed graph is decomposable. Following Giudici and Green (1999) we apply the concept of a junction tree (see Lauritzen 1996): the addition of an edge between two nodes is allowed if the two nodes belong to two different connected components of the tree or if the cliques they belong to are connected by a separator. On the other hand, the removal of an edge is allowed if such edge belongs to a single clique.

The newly sampled graph is then compared with the old graph, calculating the ratio between the two marginal likelihoods. If the ratio is greater than a predetermined threshold (acceptance probability), the proposal is accepted, otherwise it is rejected. The algorithm continues until practical convergence is reached (as in Castelo and Giudici, 2003).

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:

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

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 variance-covariance matrix Σ, and $B_0, ..., B_s$ 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:

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

We can exploit equivalences as the above one to build a dynamic graphical model, for any specified time lag in a vector autoregressive process. More formally, let $G \downarrow = \{G_1, \ldots, G_s\}$ be a collection of graphs generated at different time lags, where G_τ denotes a graph for each lag $\tau = 1, \ldots, s$.

For any given time lag τ we can then define a dynamic graphical model (DGM) as a pair (G, G_τ), where G is a graphical model, that defines the contemporaneous conditional dependences between the available random variables; and G_τ is another graphical model, that defines the temporal conditional dependences between the same variables, at two different times, lagged by τ. The vertex set and the dimension of G_τ is the same as that of G; the dimension of the edge set is also the same but the pairs on which edges can be placed are lagged: one variable at time t, and one variable at time $t - \tau$.

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

Assume that we have now available, for each time point $t = 1, \ldots, n$, a random observation from a multivariate normal distribution $\mathcal{N}_N(0, \Sigma)$.

Let $\pi^i_{t-1}$and π^i_t be the neighbors of X^i_t in the MAR and MIN networks. Following Ahelegbey et al. (2012), the marginal likelihood function decomposes according to the structure of the model into a MAR and a MIN component,

as follows:

$$P\left(X|\Sigma,G\right)=\prod_{t=1}^{n}\prod_{i=1}^{N}P\left(X_t^i|\pi_{t-1}^i(G_1),\pi_t^i(G),\Sigma\right)=$$
$$=\prod_{t=1}^{n}\prod_{i=1}^{N}P\left(X_t^i|\pi_{t-1}^i(G_1),\Sigma\right)\times\prod_{t=1}^{n}\prod_{i=1}^{N}P\left(X_t^i|\pi_t^i(G),\Sigma\right)$$

The above decomposition of the structure facilitates the inference procedure such that we can learn the MIN 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, and follow what proposed in Grzegorczyk (2010) and Ahelegby et al. (2012).

4 Empirical analysis

4.1 Data

The Bank for International Settlements locational banking statistics (LBS) include aggregate international claims and liabilities of reporting banks by country of residence and provides a plentiful data set of aggregate cross-border exposures for a set of reporting and non-reporting countries all over the world. LBS are based on quarterly data and, in practice, 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 towards all other 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 Financial networks

We first consider the application of financial network models, based on marginal correlations, to the 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 $\alpha = .05$.

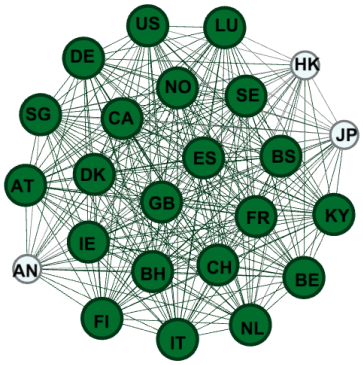

Figure 1: Marginal correlations network. The size and the intensity of the color of each node is proportional to its eigenvector centrality. Larger and darker nodes display higher centrality.

From Figure 1, note that the network is not fully connected, but only few links are removed. More precisely, most of the countries have all the 23 links with the others, with the exception of JP and HK, that have 22, and AN, that has only 21 links. 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, in accordance with Errico and Borrero (1999) and Huizinga and Nicodeme (2004).

4.3 Graphical models

We now apply graphical Gaussian model to the 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 significantly different from zero. Figure 2 shows the network obtained on the basis of the selected graphical model, taking a significance level equal to $\alpha = .05$. The size of the nodes is a function proportional to the eigenvector centrality of each node. Thicker nodes are the ones with the highest centrality, that is, most linked with respect to the other variables.

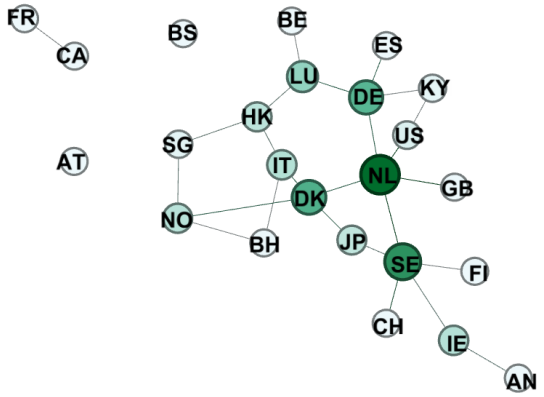

Figure 2: Partial correlation based network. The size and the intensity of the color of each node is proportional to its eigenvector centrality. Larger and darker nodes display higher centrality.

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.

From an interpretational viewpoint, note that country with the highest number of connecting edges is NL, followed by SE. 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 operate at a high cross-border level. On the other hand, the role of SE could be explained by the crisis that many northern countries suffered together at the beginning of the nineties. A second remark is that the US, as stressed by Von Peter (2007), has few connecting edges (only with KY), 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).

We remark that we have considered alternative choices for the significance level, of $\alpha = 0.10$ and $\alpha = 0.01$ and compared the obtained selected graphical models. It turns out that $\alpha = 0.10$ leads to a graph with 24 links, as for $\alpha = 0.05$, while $\alpha = 0.01$ leads to a more sparse graph, with 16 links.

4.4 Static Bayesian graphical models

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 500 iterations to calculate edge presence frequencies.

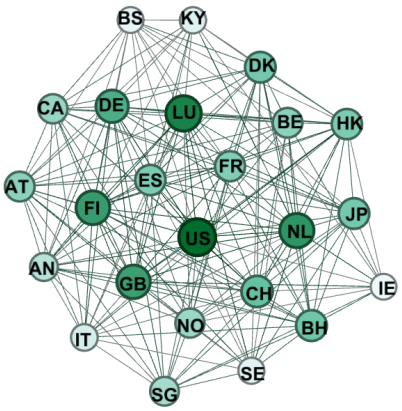

Figure 3: Static Bayesian graphical network. The size and the intensity of the color of each node is proportional to its eigenvector centrality. Larger and darker nodes display higher centrality.

From a statistical viewpoint, note that the graph in Figure 3 is somewhat "intermediate" between those in Figure 1 and Figure 2: this is the effect of Bayesian model averaging.

Economically, the results from the static Bayesian graphical model suggests that the US, followed by LU, GB, NL, FI and DE are central nodes. Indeed, the US display the highest centrality measure, followed by LU, 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, 2013), 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 prociclicity 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.

To complete the report on the static Bayesian analysis, Figure 4 contains four 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), on the cumulative difference between the number of accepted and rejected models (Figure 4c) and on the eigenvector centrality of all the 24 countries in the sample (Figure 4d).

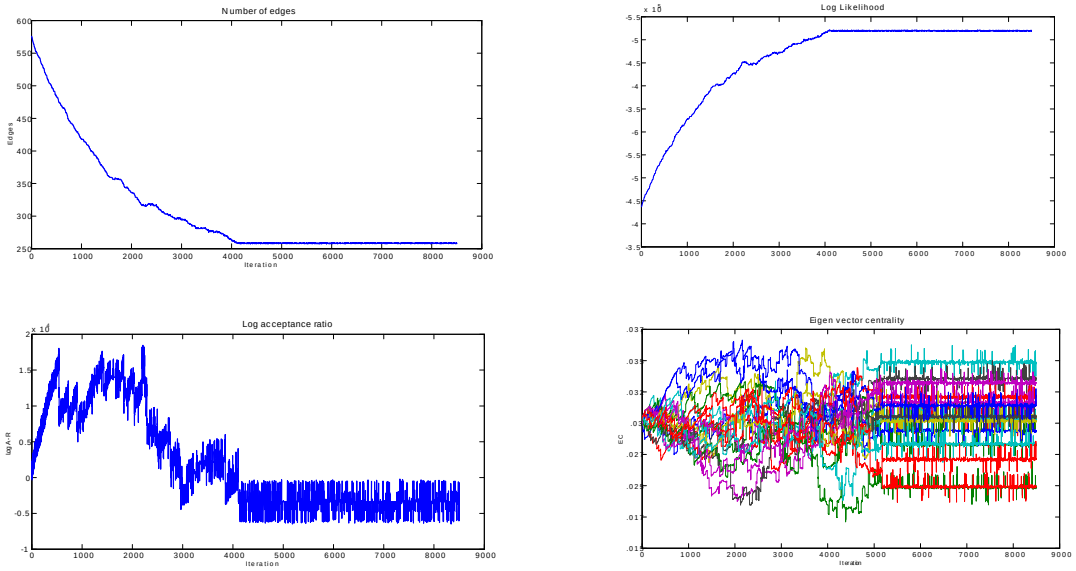

Figure 4: Convergence diagnostics for the static Bayesian graphical model. Top-left panel: number of edges. Top-right: Log-likelihood. Bottom-left:acceptance ratio. Bottom-right: eigenvector centrality

In Figure 4, and in each of the subplots, the $x-$axis represents the running iteration. Note that, after about the first 4000 iterations, the Markov chain starts to converge, according to all diagnostics.

4.5 Dynamic Bayesian graphical models

We now consider the application of dynamic Bayesian graphical models. Figure 5 shows the network based on model averaging the adjacency matrix over the different dynamic models. For the sake of comparison, we have used the same MCMC settings as before. Of course, the number of nodes is now higher and, therefore, so is the number of possible edges.

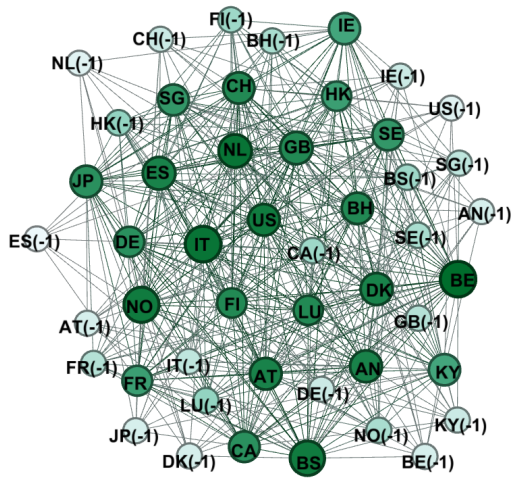

Figure 5: Dynamic Bayesian graphical network. The size and the intensity of the color of each node is proportional to its eigenvector centrality. Larger and darker nodes display higher centrality.

On the basis of Figure 5, and the associated centrality measures the countries that are most central in the MIN component of the dynamic model are BE,

IT followed by NL, US and NO. 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 from 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. In addition, there appears a lower centrality for DE, CH and the UK which are overtaken by two southern European countries: IT, ES.

As before we complete the analysis report with the diagnostics of convergence. Figure 6 shows the results.

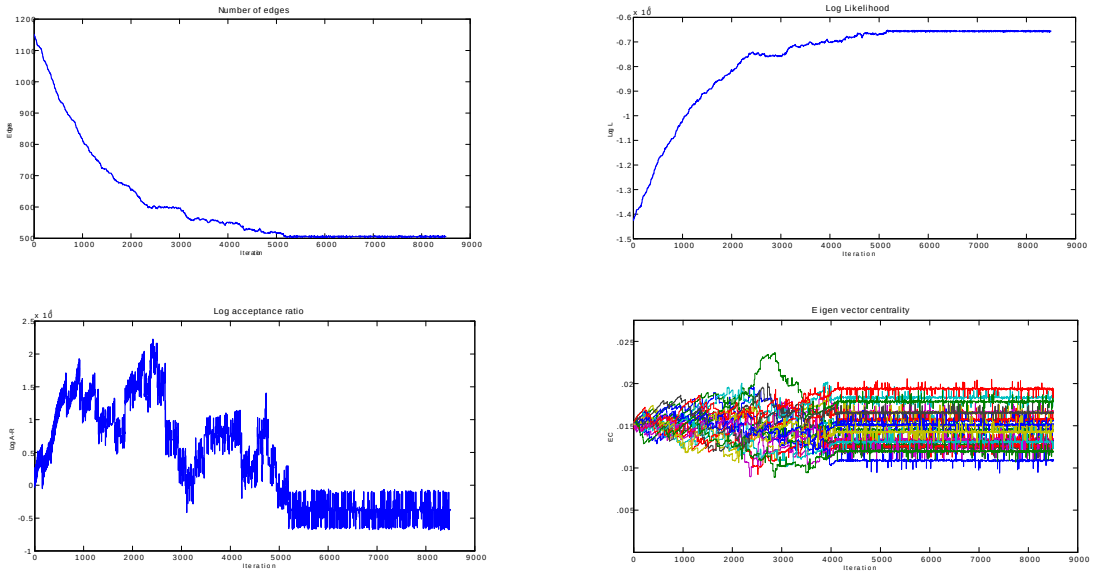

Figure 6: Convergence diagnostics for the dynamic Bayesian graphical model. Top-left panel: number of edges. Top-right: Log-likelihood. Bottom-left:acceptance ration. Bottom-right: eigenvector centrality

Figure 6 shows the actual convergence of the chosen diagnostics around 5000 iterations, for all of the plotted convergence measures.

To further check the correctness of our procedure we have applied the framework proposed by Cook et al. (2006) to both of our proposed Bayesian graphical models, considering the variance-covariance matrix, our main object of inference. For each element of Σ we have compared the values sampled by our MCMC algorithm with those obtained sampling from the prior and, then, from the conditional likelihood, as suggested in Cook et al. (2006). As summary statistics, we have calculated, for each replication, an empirical quantile for each element of the variance-covariance matrix and, then, we have computed its mean functional, from which Z scores can be obtained, leading, finally, to the Cook test statistics.

Table 3 reports some quantiles associated with the distribution of the Z-score, using the MCMC samples from our Bayesian (static) graphical model.

Prob.	0.025	0.25	.05	0.75	0.975
Quan. Stat.	0.055	0.221	0.336	0.523	0.965
Quan. Dyn.	0.022	0.165	0.320	0.490	0.923

Table 2: Quantiles associated with the cumulative probabilities of the Z-score in the static model (second row) and in the dynamic model (third row).

From Table 2 note that the distribution is concentrated around small values of the test statistics and, therefore, our procedure can be diagnosed as correct. We remark that a similar results has been obtained for the dynamic model.

We have also checked our normality assumption, employing posterior predictive analysis, as described in Geweke (2005). Our MCMC output has been summarized by two scalar functions, the determinant and the trace. For each of them we have derived the predictive density and checked whether their observed values are in a $100(1-\alpha)\%$ highest posterior density credible set. It turns out that both functions satisfy this requirement, for our Bayesian graphical models.

We finally compare our results with those obtained with Granger-causality networks (Billio et al., 2012). Although Granger causality models are causal and, therefore, cannot be strictly compared with our undirected graphical models, they are an important literature benchmark. Differently from Billio et al. (2012), that employ pairwise Granger causality, here we consider the N variable case, where $N > 2$, by estimating an n variable autoregressive model. In such conditional Granger causality, X_2 Granger-causes X_1 if lagged observations of X_2 help predict X_1 when lagged observations of all other variables $X_3 \ldots X_N$ are also taken into account.

Figure 7 shows the application of such model to our data.

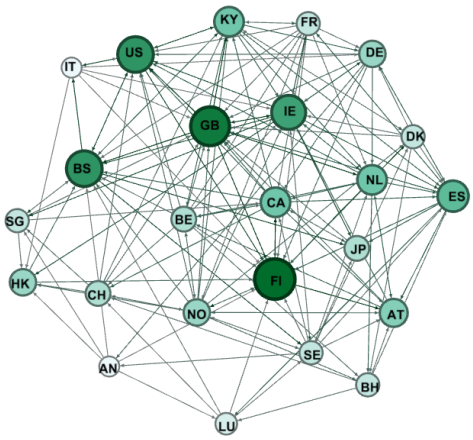

Figure 7: Conditional Granger causality network. The size and the intensity of the color of each node is proportional to its eigenvector centrality. Larger and darker nodes display higher centrality.

The network in Figure 7, and the associated centrality measures, emphasize the role of some 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 GB, and FI: this can be explained, as in the static model, by their multinational banks and, for Finland, by its local off-shore role, that generate considerable interoffice activity across borders (see also Von Peter, 2007). They are followed by BS and than by US, IE, ES and KY similarly as in the dynamic Bayesian graphical model. The novelty are the role of the Cayman Islands (KY), of Bahamas (BS) and Ireland (IE), clearly off-shore countries 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, IE is a country that also has low tax rates and lax regulation and can, therefore, be considered a "semi-tax" heaven.

We have finally evaluated a possible change in the network structure, as it can be relevant in understanding the behavior of the system and, accordingly, the systemic risk. To achieve this aim, we have split the data into two subsequent sub-samples: the first takes into account the liabilities of the countries for the period 1983-1997 and the second for the period 1997-2011. In Table 3 we report some basic network statistic for the static Bayesian model calculated for the two sub-samples, together with the same measures calculated over the whole data sample. The same statistics for the dynamic Bayesian model are shown in Table 4.

	1983-2011	1983-1997	1997-2011
Number of edges	123.5	130	124.4
Average degree	10.28	10.83	10.37
Average Eigenvector centrality	0.029	0.022	0.023
Average std of Eigenvector centrality	0.004	0.003	0.002

Table 3 : Posterior means and standard deviations for the static Bayesian graphical model

	1983-2011	1983-1997	1997-2011
Number of edges	252.5	255.5	238
Average degree	10.52	10.64	9.91
Average Eigenvector centrality	0.016	0.011	0.011
Standard deviation of Eigenvector centrality	0.008	0.004	0.004

Table 4 : Posterior means and standard deviations for the dynamic Bayesian graphical model

Table 3 shows that, for the static Bayesian graphical model, all statistics remain approximately the same over the three samples. The comparison in terms of number of edges and in terms of average degree indicates that the average number of links a node has in both networks is approximately equal. The comparison in terms of eigenvector centrality and its standard deviation leads to similar results.

Table 4 shows that the structure of the dynamic Bayesian graphical model is also quite stable over time. The most important difference between Table 4 and Table 3 is that the dynamic Bayesian network has 48 nodes (instead of only 24), and therefore, model complexity (as described, for example, by the number of edges) is higher.

The previous comparison shows that the overall structure of our Bayesian graphical models is stable in time. This does not mean that the dynamic of each single node in the network has the same property. Indeed such dynamic may be important to evaluate how systemic risk has evolved over the years, in

accordance with different stages of the world economy. In order to understand it we have therefore calculated the eigenvalue centrality of each node, for the same sample periods as before. The results are reported in Figure 8, for the static Bayesian graphical model.

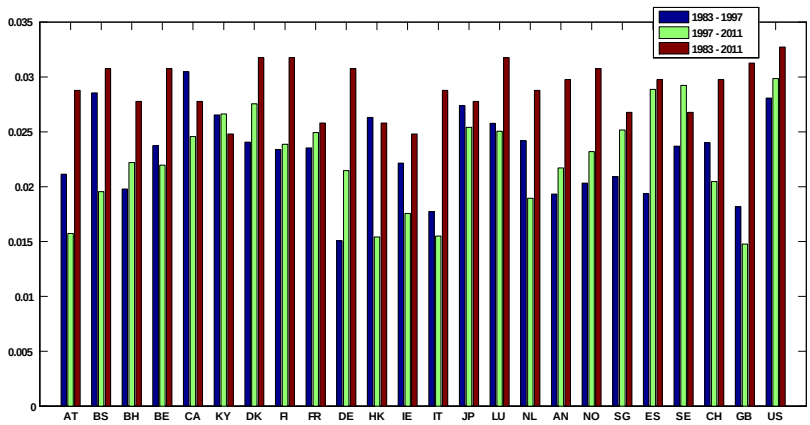

Figure 8: Eigenvector centrality for the static Bayesian model over three different peiods: (1893-1997) blue bars, (1987-2011) green bars and (1983-2011) red bas.

Figure 8 represents the eigenvector centrality scores for the static Bayesian network. For the overall sample (1983-2011), the countries with the highest centrality are US, LU, FI, DE, NL, GB, followed by CH. If we focus the attention on the first half of the sample (1983-1997), the ranking changes and the most central country turns out to be Canada, followed by US, BS, JP and two other off-shore countries, namely HK and KY. We remind that JP and HK have suffered from the Asian crisis during the mid-nineties: this fact can be the origin of their high centrality score during the first half of the sample and the relative low value in the second half of the sample. Looking instead at the second half of the sample (1997-2011) note that the US is still a high ranking central country, followed by SE, ES and DK, which do not rank as

high in the first half of the sample. We believe that such a difference could have been trigged by the European Monetary System (EMS) crises in the early nineties, which can be associated to the lost of centrality of Northern European countries.

In Figure 9 we report the eigenvector centrality scores for the nodes of the dynamic Bayesian graphical model.

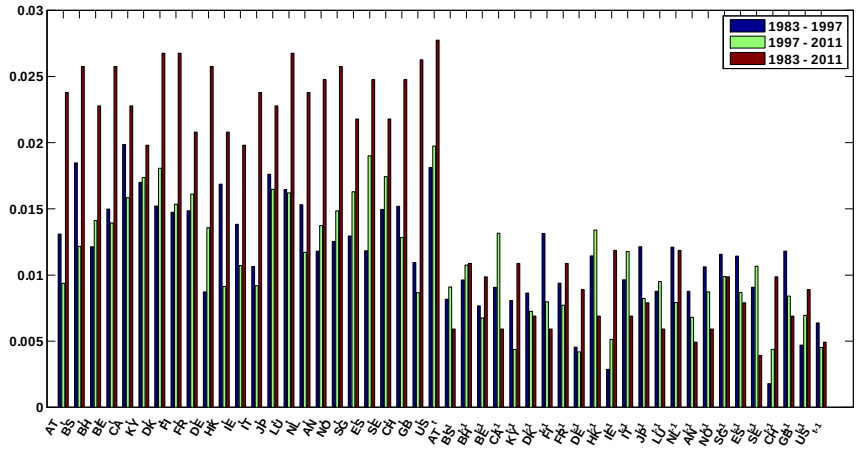

Fig. 1. Figure 9: Eigenvector centrality for the dynamic Bayesian model over three different peiods: (1893-1997) blue bars, (1987-2011) green bars and (1983-2011) red bas.

From Figure 9 it is evident that the scores associated with contemporaneous nodes are higher that the ones of the autoregressive component. The ranking of the countries in the MIR network component appears to be similar to the dynamic of the score in the static Bayesian graphical model. In the MAR component, instead, the most central countries are all European countries: DK, DE, IT, LU and CH.

Overall, Figures 8 and 9 suggest that the US maintain throughout its role as international financial hub, as its high centrality is stable over time. It is also

clear that the results obtained using the whole sample are a sort of average effect between the results obtained separately in the two sub samples. The application of the model to different subsamples, if the latter are appropriately chosen, can indeed shed further light on the causes of systemic risk: here we have highlighted the role of both the North European crisis of the early nineties and of the Asian financial crisis in the mid nineties.

5 Conclusions

Network models are a useful tool to model systemic risks in financial systems. They are essentially descriptive, and based on highly correlated networks. The paper contains two main research contributions that improve network models. First, we introduce multivariate Gaussian graphical models, defined in terms of Markov properties, on the basis of which links in the network can be removed if their corresponding partial correlation is not significant. Second, we robustify graphical model networks by means of a Bayesian approach, both in a static and in a time-varying framework, thus providing an estimate of partial correlations that, rather than being based on a single model, takes into account model uncertainty.

We have applied our proposed methods to the Bank of International Settlements locational banking statistics, with the aim of identifying central countries, whose failure could result in further distress or breakdowns in the whole system. Our results show that the countries that are potentially most contagious/subject to contagion can be splitted in three main groups: international financial hubs such as US and GB; off-shore countries such as LU, HK, BS, KY, IE and FI and, finally, countries with large cross-border financial activi-

ties 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 graphical model.

Further research include the application of what proposed here to the study of the interconnectedness between financial institutions, such as banks or insurance companies. This involves, from a methodological viewpoint, using more general models, that allow for clustering of institutions, for example within countries. On the other hand, the importance of regulatory prescriptions, based on extreme values (such as the Value at Risk) suggest the development of graphical models that model the tail, rather than the mean, of the response distribution.

6 Acknowledgement

The Authors acknowledge financial support from the Italian ministry of research grant PRIN MISURA: multivariate models for risk assessment. They also thank the referees and the associate editor for the useful comments and discussion, that have led to a considerable improvement of the paper.

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] Araujo, T. & Vilela Mendes, R. (2000). Function and form in networks of

interacting agents. Compl. Syst.12, 357–378

[3] Barigozzi, M., & Brownlees, C. (2013). NETS: network estimation for time series. Available at SSRN.

[4] 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.

[5] Boss, M., Elsinger, H., Summer, M., & Thurner, S. (2004). An empirical analysis of the network structure of the Austrian interbank market. Oesterreichesche Nationalbank's Financial stability Report, 7, 77-87.

[6] Carvalho, C.M. and West, M. (2007), Dynamic Matrix-Variate Graphical Models. Bayesian Analysis, 2(1), 69-98.

[7] Giudici, P., & Castelo, R. (2003). Improving Markov chain Monte Carlo model search for data mining. Machine learning, 50(1-2), 127-158.

[8] Cook, S. R., Gelman, A., & Rubin, D. B. (2006). Validation of software for Bayesian models using posterior quantiles. Journal of Computational and Graphical Statistics, 15(3).

[9] Corander, J., Villani, M. (2006), A Bayesian approach to modelling graphical vector autoregressions. Journal of Time Series Analysis, 27, 141-156.

[10] De Masi, G., Iori, G., & Caldarelli, G. (2006). Fitness model for the Italian interbank money market. Physical Review E, 74(6), 066112.

[11] Dawid, A.P. and Lauritzen, S.L. (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models. Annals of Statistics, 21, 1272 - 317.

[12] Edwards, D. (1990). Hierarchical interaction models. Journal of the Royal Statistical Society. Series B (Methodological), 3-20.

[13] Errico, L., & Borrero, A. M. (1999). Offshore Banking-An Analysis of Micro-and Macro-Prudential Issues. International Monetary Fund.

[14] 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.

[15] Garratt, R., Mahadeva, L., & Svirydzenka, K. (2011). Mapping systemic risk in the international banking network.

[16] Geweke, J. (2005). Contemporary Bayesian econometrics and statistics (Vol. 537). John Wiley & Sons.

[17] Giudici, P., & Green, A. P. (1999). Decomposable graphical Gaussian model determination. Biometrika, 86(4), 785-801.

[18] Grzegorczyk, M. (2010). An introduction to Gaussian Bayesian networks. In Systems Biology in Drug Discovery and Development (pp. 121-147). Humana Press.

[19] Huizinga, H., & Nicodème, G. (2004). Are international deposits tax-driven. Journal of Public Economics, 88(6), 1093-1118.

[20] 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.

[21] Lauritzen, S. L. (1996). Graphical models. Oxford University Press.

[22] Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. Journal of the American Statistical Association, 89(428), 1535-1546.

[23] Madigan, D., York, J., & Allard, D. (1995). Bayesian graphical models for discrete data. International Statistical Review/Revue Internationale de Statistique, 215-232.

[24] McGuire, P., & Tarashev, N. (2006). Tracking international bank flows. BIS Quarterly Review, 27-40.

[25] Minoiu, C., & Reyes, J. A. (2013). A network analysis of global banking: 1978–2010. Journal of Financial Stability.

[26] Pearl, J. (1988). Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann.

[27] 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.

[28] 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.

[29] Von Peter, G. (2007). International banking centres: a network perspective. BIS Quarterly Review Working Paper.

[30] Whittaker, J. (2009). Graphical models in applied multivariate statistics. Wiley Publishing.

[31] 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.