The Econometrics of Bayesian Graphical Models: A Review With Financial Application

Daniel Felix Ahelegbey

Boston University - Department of Mathematics and Statistics

28 May 2015
Research Paper

The econometrics of Bayesian graphical models: a review with financial application

Daniel Felix Ahelegbey

Department of Mathematics and Statistics, Boston University, 111 Cummington Mall, Boston, MA 02115, USA; email: dfkaye@bu.edu

(Received February 18, 2016; revised April 25, 2016; accepted May 16, 2016)

ABSTRACT

Recent advances in empirical finance have shown that the adoption of network theory is critical in order to understand contagion and systemic vulnerabilities. While interdependencies among financial markets have been widely examined, only a few studies review networks, and they do not focus on the econometrics aspects. This paper presents a state-of-the-art review on the interface between statistics and econometrics in the inference and application of Bayesian graphical models. We specifically highlight the connections and possible applications of network models in financial econometrics in the context of systemic risk.

Keywords: Bayesian inference; graphical models; model selection; systemic risk; financial crisis.

1 INTRODUCTION

The global financial crisis has prompted new research interests to understand the structure of the financial system and risk propagation channels. In order to provide a framework for strengthening financial stability, policymakers are currently not only refining the regulatory and institutional setup, but also looking for analytical tools to identify, monitor and address systemic risk better. According to Bernanke (2010) and
the Financial Crisis Inquiry Commission (2011), the global financial crisis was triggered by losses suffered by holders of subprime mortgages, and amplified by financial system vulnerabilities. They pointed out that in the absence of the systemic vulnerabilities, the triggers might produce sizable losses to certain firms, investors or asset classes, but generally they would not lead to a global recession. Many authors have found the complex connections in financial markets to have been the vulnerabilities that magnified the initial shocks of the crisis (see, for example, Acemoglu et al 2015; Battiston et al 2012; Billio et al 2012; Diebold and Yilmaz 2014; Gai et al 2011).

Network analysis has proven to be a promising tool for understanding the systemic vulnerabilities. It studies the interconnectedness among financial markets and the macroeconomy. The novelty of network analyses in this domain is crucial to model the complex connections in the financial system through information from network structure, density, homophily and centralities (Jackson 2014). While interconnectedness and contagion have been widely studied, only a few studies review networks, and they do not focus on the econometrics aspects (see Dornbusch et al (2000), Pericoli and Sbracia (2003) and Dungey et al (2005) for a review of financial contagion models). This study contributes to the literature on network econometrics. In particular, it presents a state-of-the-art review on the inference and possible applications of networks in financial econometrics in the context of systemic risk.

Many techniques for network modeling have been developed in statistics and social network literature. The common class of models includes the following: exponential random graph models (Frank and Strauss 1986; Holland and Leinhardt 1981), stochastic block models (Nowicki and Snijders 2001; Wang and Wong 1987) and latent space models (Handcock et al 2007; Hoff et al 2002). For a review of statistical models for social networks, see Kolaczyk (2009), Goldenberg et al (2010), Snijders (2011) and Kolaczyk and Csárdi (2014).

In most applications of social networks, the network is assumed to be known and is considered as the observed data. However, in the systemic risk literature, the role of interconnectedness in the risk-propagation process crucially depends on the network structure, which is generally unknown. Much of the earlier work on contagion has focused on interconnectedness arising from actual exposures among institutions, based on either balance sheet information or other financial market data. There is relatively little empirical work on the former, largely because of problems of balance sheet data accessibility. However, several studies have focused on the latter in order to understand sources of contagion and spillovers (Ahelegbey et al 2016; Barigozzi and Brownlees 2016; Billio et al 2012; Diebold and Yilmaz 2014). Although market data is readily available these days, inferring networks from observed data is characterized by uncertainty and highly complex structures.

Relationships among many real-world phenomena are often more complex than pairwise. It is well known that interring networks from data is a model determination
problem in which the number of candidates increases super-exponentially with the number of variables (Chickering et al. 2004). Therefore, applying standard techniques, such as Granger causality (Granger 1969), to identify a single model from this set of candidates often ignores the problem of model uncertainty. Also, such techniques are unable to distinguish between direct and mediated causes. Here, we present an overview of the Bayesian approach to network identification. This method takes into account network uncertainty by allowing us to incorporate prior information, where necessary, and perform model averaging (see Heckerman et al. 1995). The approach discussed in this paper is closely related to the literature on Gaussian graphical models for time series (Carvalho and West 2007; Carvalho et al. 2007; Dahlhaus and Eichler 2003). It is also related to Eichler (2007) and Zou and Feng (2009), who present the network techniques as a valid alternative to the Granger concept for causal identification and its extensions in the econometrics literature (Diks and Panchenko 2005; Hoover 2001).

We demonstrate the effectiveness of the Bayesian method in identifying interconnectedness among both the major financial sectors in the US, using the monthly returns indexes of Billio et al. (2012), and the daily volatilities of the super-sectors of the European stock market. The results on the returns network corroborate the findings of Billio et al. (2012), with evidence of higher connectedness between 2001 and 2008 and insurance companies being central to the spread of risk in the US financial market during the subperiod leading to the global financial crisis. The volatility network shows that banks and insurance companies are central to the spread of the “fear connectedness” (Diebold and Yilmaz 2014) expressed by market participants in the financial sector of the euro area.

This paper proceeds as follows. We review network applications from a statistical perspective in Section 2 and the literature on financial networks for systemic risk in Section 3. In Section 4, we relate network models to multivariate analysis and present possible applications in financial econometrics. We then discuss the Bayesian network inference in Section 5. In Section 6, we illustrate the effectiveness of the Bayesian network inference in analyzing the return and volatility connectedness of financial time series.

2 A REVIEW OF GRAPHICAL MODELS

Graphical modeling is a class of multivariate analysis that uses graphs to represent statistical models. They are represented by \((G, \theta) \in (\mathcal{G} \times \Theta)\), where \(G\) is a network, \(\theta\) is the model parameters, \(\mathcal{G}\) is the space of graphs and \(\Theta\) is the parameter space. A graph \(G = (V, E)\) is defined in terms of a set \(V\) of vertices or nodes (variables) joined by a set \(E\) of edges or links (interactions). We introduce the essential concepts of
graphical models and review the developments in statistical inference and application of Bayesian network models.

2.1 Basic terminologies

A graph with undirected edge interactions between variables is an “undirected graph” (or “Markov network”). These graphs produce a class of models commonly known as undirected graphical models, which are more suitable for analyzing similarity and correlated behaviors among variables (see Koller and Friedman 2009; Meinshausen and Bühlmann 2006; Wainwright and Jordan 2008).

A graph with directed edge interactions between variables is a “directed graph” and a directed graph without cycles is a “directed acyclic graph” (DAG). DAGs are typically based on the concept of family ordering. For instance, in A \(\rightarrow\) B \(\rightarrow\) C, A is a parent of B, and C is a child of B; A and B are ancestors of C, and B and C are descendants of A. C \(\rightarrow\) A is illegal, since a descendant cannot be his or her own ancestor. This type of ordering is suitable for expressing causal relationships. These graphs produce a class of models referred to as Bayesian networks (see, for example, Ghahramani 1998; Heckerman et al 1995; Neapolitan 2004).

A partially directed acyclic graph (or chain graph) is a type of DAG that allows bidirected edges. These graphs are suitable for applications in which a unique direction of influence cannot be ascribed to interactions among some variables. They represent a class of Markov equivalent DAGs. Two or more DAGs are said to be Markov equivalent if they depict the same set of conditional independence relationships. For example, A \(\rightarrow\) B \(\rightarrow\) C, A \(\leftarrow\) B \(\rightarrow\) C and A \(\leftarrow\) B \(\leftarrow\) C are Markov equivalent, since they all represent the conditional independence of A and C, given B (see Andersson et al 1997; Gillispie and Perlman 2001; Pearl 2000).

A bipartite graph is an undirected graph in which variables are categorized into two sets, such that nodes in one set can only interact with those in the other set, and no two nodes in the same set are connected. This type of graph belongs to the class of color graphs, in which each variable is assigned a color such that no edge connects identically colored nodes. Thus, a bipartite graph is equivalent to a two-colored graph. A factor graph is a bipartite graph with nodes categorized into variables and factors, which are represented by different shapes. Variable nodes are often represented by circles, and factor nodes are often represented by squares (see Asratian 1998; Kschischang et al 2001; Zha et al 2001).

A weighted graph is one that has a numeric value (weights) associated with each edge. A graph is complete if all vertices are connected. A clique is a subset of vertices that are completely connected. Let \(V\) be the set of vertices and \(V_A \subseteq V\); then, \(G_A\) is defined as a subgraph on nodes in \(V_A\). The triple \((V_A, V_B, V_C) \subseteq V\) forms a decomposition of a graph \(G\) if \(V = V_A \cup V_B \cup V_C\) and \(V_C = V_A \cap V_B\), such that
$G_C$ is complete and separates $G_A$ and $G_B$. The subgraph $G_C$ is called a separator. The decomposition is proper if $V_A \neq \emptyset$ and $V_B \neq \emptyset$. A sequence of subgraphs that cannot be further properly decomposed are the prime components of a graph. A graph is decomposable if it is complete, or if every prime component is complete (see Giudici and Green 1999; Koller and Friedman 2009; Wainwright and Jordan 2008).

2.2 Statistical inference

Statistical inference of the graph structure is central to the model estimation. The common methods are (i) the constraint-based approach, (ii) the score-based approach and (iii) the hybrid approach.

2.2.1 Constraint-based approach

The constraint-based approach to graph selection involves the use of statistical tests to identify conditional independence relationships among variables. The outcomes of these tests are used to constrain the graph selection process to estimate the most plausible graph that is consistent with those constraints obtained. The most widely applied constraint-based inference is the PC algorithm, based on Fisher’s $z$-transform (see Spirtes et al 2000; Verma and Pearl 1991).

2.2.2 Score-based approach

Score-based approaches are typically based on assigning a score to each candidate graph. The score represents the goodness-of-fit of the graph given the data. This approach involves a search over the set of candidates that minimizes a penalized likelihood score. Another aspect of this approach is Bayesian in nature: it usually involves priors and posterior computations, taking advantage of model averaging to address the model uncertainty problem. Examples of such algorithms are greedy search, simulated annealing and Markov chain Monte Carlo (MCMC) (Friedman and Koller 2003; Giudici and Green 1999; Madigan and York 1995).

2.2.3 Hybrid approach

The hybrid approach combines techniques from the constraint-based and score-based inferences for graph selection. Methods based on this approach are designed to adopt constraint-based reasoning as an initial step to restrict the search space for the application of the score-based scheme. An example of this is the Max-Min Hill-Climbing algorithm (Tsamardinos et al 2006). Regularization methods such as the least absolute shrinkage and selection operator (LASSO) and its variants are also hybrid methods (Banerjee et al 2008; Friedman et al 2008; Meinshausen and Bühlmann 2006; Tibshirani 1996).
2.3 Applications and developments in Bayesian network models

Graphical models have contributed to modeling challenging and complex real-world phenomena in several fields. They have been applied in forensic science as tools to aid in reasoning under uncertainty. For example, Bayesian networks have been identified as a suitable tool for analyzing evidence in complex legal and criminal cases (Dawid 2003; Wright 2007).

The study of gene interactions has become increasingly important because such information can be used as a basis for treating and diagnosing diseases, which contributes to our understanding of biological processes. Several researchers have applied graphical models to analyze gene interactions in detecting conditional dependencies (Friedman et al. 2000; Hensman et al. 2013). Several authors studied change points and time-varying interactions (see Grzegorczyk et al. 2011; Lebre et al. 2010; Robinson and Hartemink 2009).

The observation that gene data is typically characterized by heavy tails or outliers has motivated research on high-dimensional gene expressions by relaxing the assumption of normality. The active research in this area focuses on non-normality and outliers (Miyamura and Kano 2006; Vogel and Fried 2011) and non-paranormal distributions (nonparametric or semi-parametric models) to allow for mixed data (binary, ordinal or continuous) (see Teramoto et al. 2014; Zhao et al. 2012). Many studies have considered applications in which conditional distributions assume different probability models, such as Bernoulli, multinomial, Poisson and exponential families (Höfling and Tibshirani 2009; Ravikumar et al. 2010).

Studies in biological networks have revealed the complex hierarchical structures of cellular processes, which pose a challenge to researchers. An active focus in this area is designing algorithms to detect hierarchical modularity (Hao et al. 2012; Ravasz 2009), latent variables (Choi et al. 2011; Liu and Willsky 2013) and hubs (Akavia et al. 2010; Tan et al. 2014).

Graphical models have been applied in areas such as image processing, speech and handwriting recognition, which often exhibit regularities, even though they are characterized by uncertainty. Such models have been applied to acquire a structural representation of the patterns in these phenomena. Many researchers have developed state-of-the-art algorithms to track structural patterns in these domains (Bishop 2006; Koller and Friedman 2009; Murphy 2012).

Many real-world systems are too complex and complicated for humans to learn from. Machine learning has therefore become necessary to help identify the patterns in such systems. Graphical models provide a suitable framework to represent the relationships in such patterns. Developing large-scale algorithms for big data and high-dimensional problems has increasingly become a great concern in machine learning and statistics. A common approach to network inference is a centralized learning
algorithm, which is often hindered by restrictive resource constraints such as limited local computing, limited memory and expensive computational power. An active area of research in this domain is development of a decentralized system of distributed algorithms for high-dimensional problems (Liu and Ihler 2012; Meng et al 2013).

Graphical models have become more advanced in multivariate analysis, specifically with regard to multiple regression problems. To deal with high-dimensional problems, parsimony of the model is critical in achieving reasonable performances with a limited sample size. Different research directions have been considered for building a parsimonious model. Several authors have approached the problem by considering sparsity (Fan and Peng 2004; Yuan and Lin 2006). Others considered the reduced-rank approach (Bunea et al 2011; Chen et al 2013), while some investigated the sparse reduced-rank approach (Chen and Huang 2012; Lian et al 2015).

3 NETWORK ASPECT OF SYSTEMIC RISK

Systemic risk, as defined by Billio et al (2012), is “any set of circumstances that threatens the stability or public confidence in the financial system”. The European Central Bank (ECB) defines it as a risk of financial instability “so widespread that it impairs the functioning of a financial system to the point where economic growth and welfare suffer materially”. A comprehensive review of systemic risk is given in De Bandt and Hartmann (2000), Acharya et al (2010) and Brunnermeier and Oehmke (2012). Several authors have come to the same conclusion that the likelihood of major systemic crisis is related to

(1) the degree of correlation among the holdings of financial institutions,

(2) how sensitive they are to changes in market prices and economic conditions,

(3) how concentrated the risks are among those financial institutions,

(4) how closely linked they are with each other and the rest of the economy (Acharya and Richardson 2009; Brunnermeier and Pedersen 2009; Diebold and Yilmaz 2015).

Several systemic risk measures are discussed in the literature. Among them are Banking System’s (Portfolio) Multivariate Density (BSPMD; Segoviano and Goodhart 2009), conditional value-at-risk (CoVaR; Adrian and Brunnermeier 2010), absorption ratio (AR; Kritzman et al 2011), marginal expected shortfall (MES; Acharya et al 2010; Brownlees and Engle 2011), distressed insurance premium (DIP; Huang et al 2012), dynamic causality index (DCI) and principal component analysis systemic (PCAS) risk measures (Billio et al 2012) and network connectedness measures (NCMs; Diebold and Yilmaz 2014). The BSPMD embeds banks’ distress
interdependence structures, which capture distress dependencies among the banks in the system; CoVaR measures the value-at-risk (VaR) of the financial system, conditional on an institution being under financial distress; AR measures the fraction of the total variance of a set of \( N \) financial institutions explained or “absorbed” by a finite number \( K < N \) of eigenvectors; MES measures the exposure of each individual firm to shocks of the aggregate system; DIP measures the insurance premium required to cover distressed losses in the banking system; the DCI captures how interconnected a set of financial institutions is by computing the fraction of significant Granger causality relationships among their returns; PCAS captures the contribution of an institution to the multivariate tail dynamics of the system; and NCMs aggregate the contribution of each variable to the forecast error variance of other variables across multiple return series.

Systemic risks are the major contributors to financial crises. Since the 1990s, it has been observed that financial crises appear in clusters (for instance, the East Asian crisis in 1997, the Russia and Long-Term Capital Management (LTCM) crisis in 1998, the Brazil crisis in 1999, the dot-com crisis in 2000, the Argentina crisis in 2001, the Iceland and Turkey crises in 2006, the China crisis in 2007, the global financial crisis in 2007–9 and the European crisis in 2010–13). As to whether systemic risks can be reliably identified in advance, Bernanke (2013) analyzed the global financial crisis by distinguishing between triggers and vulnerabilities of the system that caused the event. This distinction is helpful and allows us to identify which factors to focus on to guard against a repetition of the global financial crisis. The triggers are the events that began the crisis. One prominent trigger of the global financial crisis were the losses suffered by holders of subprime mortgages. Tang et al (2010) discussed how triggers for crises differ, and examples include sovereign debt default, risk management strategies, sudden stops in capital flows, collapses of speculative bubbles, inconsistencies between fundamentals and policy settings and a liquidity squeeze. Billio et al (2012) found liquidity and credit problems to be the triggers of both the LTCM crisis and the global financial crisis.

The vulnerabilities are the preexisting structural weaknesses of the financial system that amplified the initial shocks (Bernanke 2013). Examples of such factors include a lack of macroprudential focus in regulation and supervision, high levels of leverage and complex interconnectedness. In the absence of these vulnerabilities, the triggers might produce sizable losses to certain firms, investors or asset classes, but generally they would not lead to a global recession. Many authors have found complex financial interdependencies to be channels that magnify initial shocks to the system (Acemoglu et al 2015; Billio et al 2012; DasGupta and Kaligounder 2014; Diebold and Yilmaz 2014, 2015; Gai et al 2011; Tang et al 2010). Tang et al (2010) showed that financial crises are indeed alike, as all linkages are statistically important across all crises. They identified three potential channels for contagion effects:
(1) idiosyncratic channels, which provide a direct link from the source asset market to international asset markets;

(2) market channels, which operate through either the bond or stock markets;

(3) country channels, which operate through the asset markets of a country jointly.

Billio et al (2012) found that the 2007–9 period experienced a higher level of interconnectedness and systemic vulnerability than the LTCM. This allowed the authors to explain why the impact of the global financial crisis affected a much broader segment of financial markets and threatened the viability of several important financial institutions to a greater extent than the LTCM. Acemoglu et al (2015) showed that, beyond a certain point, dense interconnections serve as a mechanism for the propagation of shocks, leading to a more fragile financial system. Diebold and Yilmaz (2015) showed that the impact of systemic risk depends on the collective action and connectedness of financial institutions as well as interaction between financial markets and the macroeconomy.

Bernanke (2013) argued that shocks are inevitable, but identifying and addressing vulnerabilities is key to ensuring a robust financial system. To understand vulnerabilities, researchers and regulators are currently focusing on network analyses to identify complex connectivities in financial markets. Some of the applications of network tools include measuring the degree of connectivity of particular financial institutions to determine systemic importance, forecasting the likely contagion channels of institutional default or distress and visualizing the “risk map” of exposure concentrations and imbalances in the system (Bisias et al 2012). To ensure a robust financial system, it is of crucial importance to

(a) identify systemically important institutions,

(b) identify specific structural aspects of the system that are particularly vulnerable,

(c) identify potential mechanisms for shock propagation in the system.

Early studies on systemic risk networks focused on linkages arising from actual exposures based on balance sheet information. However, due to data accessibility issues, little empirical work has been done in this area (Cont et al 2013; Georg 2013). Meanwhile, several statistical and econometric methods have been advanced to study interdependencies, contagion and spillover effects from observed market data. The commonly discussed approaches in recent literature include the following: Granger causality (Billio et al 2012), variance decomposition (Diebold and Yilmaz 2014), tail risk (Hautsch et al 2015) and partial correlation (Barigozzi and Brownlees 2016). Most of these approaches often ignore the problem of network uncertainty and highly complex structures. In this paper, we discuss a Bayesian approach to network inference that helps to address the above problems.
4 GRAPHICAL MODELS IN MULTIVARIATE ANALYSIS

Graphical models have become popular for modeling patterns in complex systems due to their ability to provide an intuitive interpretation of the interactions. The graphs present a way of visualizing the relationships between variables in order to distinguish between direct and indirect interactions. The idea of connecting the multivariate time series literature and graphical models is gradually becoming a vibrant field of research in economics and finance. We relate graphical models to multivariate statistical analysis, specifically in multivariate regression problems, with possible applications in financial econometrics.

4.1 Multivariate (multiple) regression

A typical multivariate multiple regression model is given by

\[ Y = BX + U, \]

where \( X \) and \( Y \) are vectors of exogenous and response variables, respectively, \( B \) is a coefficient matrix, and \( U \) is a vector of errors. The common approach in much empirical research is to fit the above model and test for restrictions. In testing for the statistical significance of each of the estimated coefficients, we typically specify an acceptable maximum probability of rejecting the null hypothesis when it is true, i.e., committing a type I error. In multiple hypothesis testing, the type I errors committed increase with the number of hypotheses, which may have serious consequences for the conclusions and generalization of the results. Several approaches have been proposed to deal with this problem: see Shaffer (1995) and Drton and Perlman (2007) for a review and discussion of multiple hypothesis testing.

Graphical models provide a convenient framework for exploring multivariate dependence patterns. By considering (4.1) as a causal (dependency) pattern of elements in \( Y \) on elements in \( X \), the coefficient matrix under a null hypothesis of single restrictions is \( B_{ij} = 0 \) if \( y_i \) does not depend on \( x_j \), and \( B_{ij} \neq 0 \) otherwise. We define a binary connectivity matrix, \( G \), such that \( G_{ij} = 1 \) implies \( x_j \rightarrow y_i \), \( \iff \) \( B_{ij} \neq 0 \).

Thus, \( G \) can be interpreted as a (directed) graph of the conditional dependencies between elements in \( X \) and \( Y \). \( B \) can be represented as

\[ B = (G \circ \Phi), \]

where \( \Phi \) is a coefficient matrix and the operator \( (\circ) \) is the element-by-element Hadamard product (i.e., \( B_{ij} = G_{ij} \Phi_{ij} \)). There is a one-to-one correspondence between \( B \) and \( \Phi \) conditional on \( G \), such that \( B_{ij} = \Phi_{ij} \) if \( G_{ij} = 1 \), and \( B_{ij} = 0 \) if \( G_{ij} = 0 \). Thus, we interpret \( \Phi \) (\( B \)) as the unconstrained (respectively, constrained) regression.
coefficient matrix. For example,

\[
y_1 = 1.3x_1 + 0.5x_3 + u_1 \\
y_2 = 0.9x_1 + 0.5x_5 + u_2 \\
y_3 = x_2 + 0.7x_4 + u_3
\]

\[
B = \begin{pmatrix}
  x_1 & x_2 & x_3 & x_4 & x_5 \\
  y_1 & 1.3 & 0 & 0.5 & 0 & 0 \\
  y_2 & 0.9 & 0 & 0 & 0 & 0.5 \\
  y_3 & 0 & 1 & 0 & 0.7 & 0
\end{pmatrix}
\]

\[
(4.3)
\]

\[
G = \begin{pmatrix}
  y_1 & 1 & 0 & 1 & 0 & 0 \\
  y_2 & 1 & 0 & 0 & 0 & 1 \\
  y_3 & 0 & 1 & 0 & 1 & 0
\end{pmatrix}
\]

\[
\Phi = \begin{pmatrix}
  x_1 & x_2 & x_3 & x_4 & x_5 \\
  y_1 & 1.3 & a_1 & 0.5 & a_2 & a_3 \\
  y_2 & 0.9 & a_4 & a_5 & a_6 & 0.5 \\
  y_3 & a_7 & 1 & a_8 & 0.7 & a_9
\end{pmatrix}
\]

\[
(4.4)
\]

where \(a_i \in \mathbb{R}, i = 1, \ldots, 9\), are expected to be statistically not different from zero by definition. The expression (4.4) presents an interesting alternative for modeling (4.3). Instead of estimating the unconstrained coefficient matrix, \(\Phi\), and performing multiple hypothesis tests, a more efficient alternative is to infer \(G\) as a variable selection matrix to estimate only the relevant coefficients in \(B\). Thus, nonzero elements in \(B\) correspond to nonzero elements in \(G\). Inference of \(G\), taking into account all possible dependence configurations, automatically handles the multiple testing problem in multivariate multiple regression models.

In most regression models, the parameters to be determined are \(\{B, \Sigma_u\}\), where \(\Sigma_u\) is the covariance matrix of \(U\). In graphical models, inference of the graph structure is central to the model estimation, and the set of parameters, \(\theta\), describes the strength of the dependence among variables. Hence, by relating graphical models to multivariate regressions, \(\theta\) must be equivalent to the regression parameters, ie, \(\theta = \{B, \Sigma_u\}\).

Let \(Y\) and \(X\) denote an \(n_y \times 1\) and an \(n_x \times 1\) vector of dependent and explanatory variables, respectively. Let \(Z = (Y', X')'\) be the \(n = (n_y + n_x) \times 1\) vector of stacked \(Y\) and \(X\). Suppose the joint distribution, \(Z\), follows the distribution \(Z \sim \mathcal{N}_n(0, \Omega^{-1})\), where \(\Sigma = \Omega^{-1}\) is an \(n \times n\) covariance matrix. The joint distribution of \(Z\) can be summarized with a graphical model, \((G, \theta)\), where \(G\) is of dimension \(n_y \times n_x\) and consists of directed edges from elements in \(X\) to elements in \(Y\). Therefore, estimating the model parameters associated with \(G\) is equivalent to estimating \(\Omega\), ie, \(\theta = \Omega\). Given \(\Omega\), the parameters of model (4.1) can be found from \(\Sigma = \Omega^{-1}\) as

\[
B = \Sigma_{yx} \Sigma_{xx}^{-1} \quad \Sigma_e = \Sigma_{yy} - \Sigma_{yx} \Sigma_{xx}^{-1} \Sigma_{xy},
\]

\[
(4.5)
\]

where \(\Sigma_{xy}\) is \(n_x \times n_y\) covariances between \(X\) and \(Y\), \(\Sigma_{yy}\) is \(n_y \times n_y\) covariances among \(Y\) and \(\Sigma_{xx}\) is \(n_x \times n_x\) covariances among \(X\). For further computational
aspects of graphical models, see Heckerman and Geiger (1994) and Lenkoski and Dobra (2011).

4.2 Applications in econometrics and finance

We now present potential applications of graphical models in econometrics and finance.

4.2.1 Structural model estimation

Vector autoregressive (VAR) models are widely used to estimate and forecast multivariate time series in macroeconomics. It is generally known that such models do not have direct economic interpretations. However, due to their ability to forecast dynamics in macroeconomic variables, such a limitation is overlooked. Structural VARs (SVARs), however, have direct economic interpretations, but these are not estimable due to identification issues. For SVAR identification, the standard approach relies on a reduced-form VAR to determine the relationships among shocks as a means of providing economic intuition about the structural dynamics. To achieve this, some researchers impose structures provided by a specific economic model, in which case the empirical results will only be as credible as the underlying theory (Kilian 2013). Moreover, in many cases, there are not enough credible exclusion restrictions to achieve identification.

Following (4.2), the SVAR model can be expressed in a graphical model form as

\[
Y_t = \sum_{i=0}^{p} B_i Y_{t-i} + \varepsilon_t = \sum_{i=0}^{p} (G_i \circ \Phi_i) Y_{t-i} + \varepsilon_t,
\]  

(4.6)

where \((G_0 \circ \Phi_0)\) and \((G_s \circ \Phi_s)\), \(s \geq 1\), are the graphical models representing the cross-sectional and temporal dependences, respectively.\(^1\)

4.2.2 Time-varying model estimation

Fixed or time-varying parameter models are standard applications in most empirical works. These approaches to modeling real-world phenomena implicitly assume that interactions between variables are stable over time, and only the parameters vary or are fixed. This assumption may have consequences for the performance of estimated models. Some empirical works have shown that financial networks especially exhibit

random fluctuations over various time scales (Billio et al 2012), which must be incorporated into modeling dynamics in observed data. A typical time-varying model of (4.1) is given by

\[ Y_t = B_t X_{t-1} + U_t. \]  

Following the expression in (4.2), the coefficient matrix of (4.7) can be expressed as

\[ B_t = G \circ \Phi_t, \quad B_t = G_t \circ \Phi, \quad B_t = G_t \circ \Phi_t. \]  

The first expression of (4.8) follows the typical time-varying parameter models commonly discussed in most empirical papers, where the graph is invariant over time. The other two expressions (4.8) present cases of time-varying structures of dependence. The state of the world is not constant over time, so devoting our attention to modeling dynamics of the structure of interaction seems to be a more interesting line of research in order to understand the ever-transforming modern economic and financial system.2

4.2.3 High-dimensional model estimation

There is an increasing interest in high-dimensional models and big data analysis. This has become necessary, as many studies have shown that information from large data sets enriches existing models, produces better forecasts for VAR models and also reveals the connectedness of the financial system. Graphical models are therefore relevant for high-dimensional modeling, as they offer interpretations of information extracted from large data sets.3

The graphical approach can be used to build models that serve as alternatives to the factor methods when dealing with large data sets. In factor-augmented VAR models, information is extracted from a large number of variables to build factors to augment the VAR. Following the justification of this approach (Bernanke et al 2005), graph search algorithms can be applied to select relevant predictors from a large set of exogenous variables, which can be used to augment the VAR. This method will provide a more interpretable model than the factor approach.

4.2.4 CAPM-like model estimation

A fundamental model in financial theory is the capital asset pricing model (CAPM; Sharpe 1964). This is an extension of the portfolio theory of Markowitz (1952). This

---

2 See Bianchi et al (2014) for an example of a graphical factor model with Markov-switching graphs and parameters for modeling contagion. See also Carvalho and West (2007) for graphical multivariate volatility modeling induced by time-varying covariances across series.

3 See Ahelegbey et al (2014) for a discussion on modeling sparsity in large graphical VAR models with uncertainty on the lag order. See also Jones et al (2005) and Scott and Carvalho (2008) for discussions on approaches for penalizing globally or locally “dense” graphs when estimating high-dimensional models.
approach has received much criticism due to problems of empirical evidence. Fama and French (2004) summarized the popularity of the CAPM as follows:

The CAPM’s empirical problems may reflect theoretical failings, the result of many simplifying assumptions. But they may also be caused by difficulties in implementing valid tests of the model.

Graphical models can be applied to decompose asset return correlations into market specific and idiosyncratic effects, as in the classical CAPM models.4

4.2.5 Portfolio selection problem

Portfolio risk analysis is typically based on the assumption that the securities in the portfolio are well diversified. A well-diversified portfolio is one that is exposed only to market risk within asset classes and includes a variety of significantly different asset classes. Precisely, a well-diversified portfolio is made up of asset classes that are highly uncorrelated and are considered to be complementary. The lesser the degree of correlation, the higher the degree of diversification, and the lower the number of asset classes required. Portfolios that contain securities with several correlated risk factors do not meet this diversification criteria. Despite unprecedented access to information, some portfolios, by construction, contain a predominant factor, and most risk-modeling techniques are unable to capture this contagion.

To measure diversification more accurately, the graphical approach can be applied to study the structure of interconnection among the asset classes. Since the arbitrage pricing theory (APT) model of Roll (1977) and Ross (1976) and the return-based style model of Sharpe (1992) are regression models, the graphical approach provides a useful technique for portfolio selection by explicitly modeling the dependence of the factors or asset classes.5

4.2.6 Risk-management-style assessment

Sharpe (1992) introduced a return-based analysis to measure management style and performance. The analysis is based on the idea that a manager builds a portfolio according to a specific investment philosophy, and investments reflect a style. The approach is based on a style-regression model to determine the “effective mix”, given by the estimated model, which represents the return from the style, while the residuals reflect the performance due to the “selection” (active management).

---

4 See Ahelegbey and Giudici (2014) for a discussion on Bayesian hierarchical graphical models, which allow correlations to be decomposed into a country (market) effect plus a bank-specific (idiosyncratic) effect.

Suppose we are interested in modeling the annual return on an investment. This will depend on factors such as the amount invested ($Y$), the annual stock market increase ($S$) and the experience or style of the fund manager ($E$). A simple network to model this interaction is displayed in Figure 1, where $R = f(E, S, Y)$, and the returns are a function of $Y$, $S$ and $E$.

This can serve as a benchmark to assess the performance of fund managers. For instance, since the fund manager’s level of experience may be unknown, a qualitative measure can be applied to rank their experience, and a probabilistic inference can be obtained on the likely level of experience, given information on the other variables in the network.

5 BAYESIAN INFERENCE PROCEDURE FOR STRUCTURE LEARNING

Network structure learning using standard frequentist techniques presents multiple testing problems. It also involves exploring all candidate structures of the model, which poses a challenge since the space of possible structures increases super-exponentially (Chickering et al 2004). The standard practice of identifying a single model that summarizes these relationships often ignores the uncertainty problem. The Bayesian methodology has proved to be more efficient in addressing uncertainties and complexities in the inference problem than standard frequentist techniques (Heckerman et al 1995). The Bayesian approach to network modeling is a class of probabilistic graphical models, where each node represents a random variable, and the links express probabilistic relationships between these variables. The network captures the way in which the joint distribution over all the random variables can be decomposed into a product of factors, each depending only on a subset of the variables.

6 See Ammann and Verhofen (2007) for a network application for mutual fund managers behavior analysis.
The description of the Bayesian inference is completed with the elicitation of prior distributions on the graphs and parameters, the posterior approximations and graph estimation.

## 5.1 Prior distribution

Modeling the joint distribution of \((G, \Omega)\) from a Bayesian perspective can be expressed in a natural hierarchical structure \(P(G, \Omega) = P(G)P(\Omega \mid G)\). In the absence of genuine prior knowledge of the network of interactions, the common approach is to assume a uniform prior for \(G\), i.e., \(P(G) \propto 1\). (For discussions on other graph priors, see Friedman and Koller (2003), Jones et al (2005) and Scott and Carvalho (2008).)

The standard parameter prior for graphical models is often conditioned on the graphs. Two main classes of such priors are commonly discussed in the literature. The first is based on DAG models that permit an unconstrained precision matrix \(\Omega\) (see Consonni and Rocca 2012; Geiger and Heckerman 2002; Grzegorczyk 2010; Heckerman et al 1995). The second is based on decomposable undirected graph (UG) models, which allow constraints or no constraints on \(\Omega\) (see Carvalho and Scott 2009; Roverato 2002; Wang and Li 2012). An unconstrained \(\Omega\) often characterizes a complete graph with no missing edges. The standard parameter prior for Gaussian DAG models with zero expectations is a Wishart distribution, whereas that of UG models is a G-Wishart or hyper-inverse Wishart distribution.

We follow the standard parameter prior for Gaussian DAG models, with density given by

\[
P(\Omega \mid G) = \frac{1}{K_n(v, S)} |\Omega|^{(v-n-1)/2} \exp\{-\frac{1}{2} \{\Omega, S\}\}, \tag{5.1}
\]

where \(\langle A, B \rangle = \operatorname{tr}(A'B)\) denotes the trace inner product, \(v > n + 1\) is the degrees of freedom parameter, \(S\) is a prior sum of squares matrix and \(K_n(v, S)\) is the normalizing constant:

\[
K_n(v, S) = 2^{vn/2} |S|^{-v/2} \Gamma_n \left(\frac{v}{2}\right),
\]

\[
\Gamma_b(a) = \pi^{b(b-1)/4} \prod_{i=1}^{b} \Gamma\left(a + \frac{1-i}{2}\right), \tag{5.2}
\]

where \(\Gamma_b(a)\) is the multivariate generalization of the Gamma function, \(\Gamma(\cdot)\). In this application, we set \(v = n + 2\) and \(S = vI_n\), where \(I_n\) is the \(n\)-dimensional identity matrix.
5.2 Posterior approximation

Let $Z_t = (Y_t', X_t')'$, where $Y_t$ represents the set of dependent variables at time $t$ and $X_t$ is the set of explanatory variables. Suppose that $Z_t \sim \mathcal{N}_n(0, \Omega^{-1})$, where $n$ is the number of elements in $Z_t$. Let $\mathcal{D} = (Z_1, \ldots, Z_T)$ be a complete series of the observed variables. Then, the likelihood function $P(\mathcal{D} \mid \Omega, G)$ is multivariate Gaussian with density

$$P(\mathcal{D} \mid \Omega, G) = (2\pi)^{-nT/2} |\Omega|^{T/2} \exp\left\{-\frac{1}{2} \{\Omega, \hat{S}\}\right\},$$

(5.3)

where $\hat{S} = \sum_{t=1}^T Z_tZ_t'$ is the $n \times n$ sum of squares matrix. Since $G$ is unknown, we estimate the graph by integrating out $\Omega$ analytically to produce a marginal likelihood function

$$P(\mathcal{D} \mid G) = \int P(\mathcal{D} \mid \Omega, G) P(\Omega \mid G) \, d\Omega = (2\pi)^{-nT/2} \frac{K_n(v + T, \bar{S})}{K_n(v, \overline{\bar{S}})},$$

(5.4)

where $\overline{\bar{S}} = \overline{S} + \hat{S}$ is the posterior sum of squares matrix. The graph posterior is given by $P(G \mid \mathcal{D}) = P(G) P(\mathcal{D} \mid G)$. Based on the uniform prior assumption over $G$, maximizing $P(G \mid \mathcal{D})$ is equivalent to maximizing $P(\mathcal{D} \mid G)$. Following the standard Bayesian paradigm, (5.4) can be factorized into a product of local terms, each involving a response variable ($y_i$) and its set of selected predictors ($\pi_i$):

$$P(\mathcal{D} \mid G) = \prod_{i=1}^{n_y} P(\mathcal{D} \mid G(y_i, \pi_i)) = \prod_{i=1}^{n_y} \frac{P(\mathcal{D}(y_i, \pi_i) \mid G)}{P(\mathcal{D}(\pi_i) \mid G)},$$

(5.5)

where $\pi_i = \{j = 1, \ldots, n_x: G_{ij} = 1\}$, $n_y$ is the number of equations and $G(y_i, \pi_i)$ is the subgraph of $G$, with links from $\pi_i$ to $y_i$. $\mathcal{D}(y_i, \pi_i)$ and $\mathcal{D}(\pi_i)$ are submatrixes of $\mathcal{D}$, consisting of ($y_i, \pi_i$) and $\pi_i$, respectively. The closed form of (5.5) is

$$P(\mathcal{D}^k \mid G) = (\pi)^{-n_kT/2} \frac{|S_{k,k}|^{v/2}}{|S_{k,k}|^{(v+T)/2}} \prod_{i=1}^{n_k} \frac{\Gamma((v + T + 1 - i)/2)}{\Gamma((v + 1 - i)/2)},$$

(5.6)

where $k \in \{(y_i, \pi_i), \pi_i\}$ is of dimension $n_k$, $\mathcal{D}^k$ is a submatrix of $\mathcal{D}$ associated with $k$, and $|S_{k,k}|$ and $|\overline{\bar{S}}_{k,k}|$ are the determinants of the prior and posterior sum of the square matrices of $\mathcal{D}^k$.

5.3 Graph estimation

We sample the graph following the MCMC algorithm in Madigan and York (1995). The scheme is such that, at the $r$th iteration, given $G^{(r-1)}$, the sampler proposes a new graph $G^{(*)}$ with acceptance probability

$$A(G^{(*)} \mid G^{(r-1)}) = \min\left\{ \frac{P(\mathcal{D} \mid G^{(*)}) P(G^{(*)}) Q(G^{(r-1)} \mid G^{(*)})}{P(\mathcal{D} \mid G^{(r-1)}) P(G^{(r-1)}) Q(G^{(*)} \mid G^{(r-1)})}, 1 \right\},$$

(5.7)
where \( Q(G^{(e)} | G^{(r-1)}) \) and \( Q(G^{(r-1)} | G^{(e)}) \) are the forward and reverse proposal distribution, respectively. The proposal distribution assigns a uniform probability to all possible edges that can be reached from the current state \( G^{(r-1)} \) by adding or deleting a single edge. If the new graph \( G^{(e)} \) is accepted, then the graph at the \( r \)th iteration is set to \( G^{(r)} = G^{(e)} \); otherwise, \( G^{(r)} = G^{(r-1)} \).}

6 FINANCIAL APPLICATION

We now illustrate the effectiveness of the Bayesian approach to network inference in analyzing return and volatility connectedness of financial time series.

6.1 Financial system interconnectedness

We study the structure of interconnectedness among the four major sectors of the US financial system, following Billio et al (2012). The data consists of monthly return indexes for hedge funds (HF), banks (BK), brokers (BR) and insurance companies (IN) in the United States from January 1994 to December 2008. Single hedge-fund data was obtained from the Lipper TASS database. Data for individual banks, brokers and insurers was obtained from the Center for Research in Security Prices database. The monthly returns of all companies with standard industrial classification (SIC) codes 6000–6199, 6200–6299 and 6300–6499 were used to construct value-weighted indexes for banks, brokers and insurers, respectively.

We estimate the temporal dependence pattern in a VaR with lag order \( (p = 1) \), chosen according to a Bayesian information criterion (BIC). We examine the network for two sample periods (1994–2000 and 2001–8). Table 1 shows the posterior probabilities of the presence of edges in the network. The left (respectively, right) panel shows the edge posterior probabilities in the network for the period 1994–2000 (respectively, 2001–8). Bold values indicate links for which posterior probabilities are greater than 0.5 under a 95% credibility interval. The edges are directed from column labels (at \( t_1 \)) to row labels (at \( t \)). The edge posterior probabilities in the first period (1994–2000) are very low compared with the second (2001–8). Thus, we find no evidence of significant linkages in the first period. However, in the second period, we find strong interconnectedness among the institutions. More specifically, we find evidence of an autoregressive effect among hedge funds, ie, \( P(\text{HF}_{t-1} \rightarrow \text{HF}_t \mid D) = 0.95 \), and a strong effect of insurers on brokers, \( P(\text{IN}_{t-1} \rightarrow \text{BR}_t \mid D) = 0.98 \); insurers on banks, \( P(\text{IN}_{t-1} \rightarrow \text{BK}_t \mid D) = 0.82 \); insurers on hedge funds, \( P(\text{IN}_{t-1} \rightarrow \text{HF}_t \mid D) = 0.71 \); banks on brokers, \( P(\text{BK}_{t-1} \rightarrow \text{BR}_t \mid D) = 0.75 \); and brokers on insurers, \( P(\text{BR}_{t-1} \rightarrow \text{IN}_t \mid D) = 0.64 \).

7 See Ahelegbey et al (2016) for discussions on the convergence diagnostics of the MCMC, estimation of the edge posterior probabilities and a pseudo-code of the algorithm.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \text{HF}_{t-1} )</td>
<td>( \text{BR}_{t-1} )</td>
</tr>
<tr>
<td>( \text{HF}_t )</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td>( \text{BR}_t )</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>( \text{BK}_t )</td>
<td>0.26</td>
<td>0.20</td>
</tr>
<tr>
<td>( \text{IN}_t )</td>
<td>0.18</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Bold values indicate links with probabilities greater than 0.5 under a 95% credibility interval.

FIGURE 2  Network of hedge funds (HF), brokers (BR), banks (BK) and insurance (IN) between (a) 1994–2000 and (b) 2001–8, without self-loops.

The blue (red) links are lagged positive (negative) effects.

The network representation of the results of Table 1 is shown in Figure 2. This figure reveals the direction and sign (represented by the colored arrows) of the linkages among the institutions. The blue (respectively, red) links are lagged positive (respectively, negative) effects. We see no links in the first period. In the second period, we find a negative effect of banks on brokers, a bidirectional positive link between insurers and brokers and a positive effect of insurers on banks and hedge funds. These results corroborate the findings of Billio et al (2012), providing evidence of a higher vulnerability in the system between 2001 and 2008. We also find that insurers are central to the spread of risk in the system during the 2001–8 period.

6.2 Volatility connectedness in the euro area

Volatility networks (also referred to as “fear connectedness”) have become increasingly important due to their ability to track the fear of investors and identify risk transmission mechanisms in the financial system (Diebold and Yilmaz 2014).
analyze the “fear connectedness” in the European stock market using intraday high–low price indexes of the nineteen super-sectors of EURO STOXX 600, obtained from Datastream and covering the period September 1, 2006 to September 19, 2014. See Table 2 for a list of the super-sectors that represent the largest euro area companies by the Industry Classification Benchmark (ICB). The institutions cover countries such as Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain.

Let \( p_{h,i,t} \) and \( p_{l,i,t} \) denote the highest and lowest prices of stock \( i \) on day \( t \). We obtain the intraday price range as a measure of the realized volatilities

\[
RV_{i,t} = \frac{1}{4\log(2)}(\log p_{h,i,t} - \log p_{l,i,t})^2.
\]

We study the connectedness of the log volatilities as the dependence pattern in a VaR(1). We characterize the dynamics of the connectedness using a rolling estimation with a window size of 100 days. We compare the Granger causality network (henceforth GCnet) with that of the Bayesian method (henceforth BGnet).

We present in Figure 3 the dynamics of the total connectedness index of the GCnet (in blue) and the BGnet (in red) with their respective BIC scores over the sample period. We notice a significant difference in the total number of interconnections among the institutions (see Figure 3(a)). We also see that the GCnet records higher linkages than the BGnet. This is not surprising, since the Granger causality approach deals only with bivariate time series and is unable to distinguish between direct and mediated causal influences in multivariate settings. The Bayesian method, however, considers single and multiple testing possibilities, and it is therefore able to identify direct and mediated causal effects. The BIC score (see Figure 3(b)) shows that the BGnet produces relatively better structures than the GCnet.

### Table 2 Description of EURO STOXX 600 super-sectors.

<table>
<thead>
<tr>
<th>No</th>
<th>Name</th>
<th>ID</th>
<th>No</th>
<th>Name</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Banks*</td>
<td>BK</td>
<td>11</td>
<td>Media</td>
<td>MD</td>
</tr>
<tr>
<td>2</td>
<td>Insurance companies*</td>
<td>IN</td>
<td>12</td>
<td>Travel &amp; leisure</td>
<td>TL</td>
</tr>
<tr>
<td>3</td>
<td>Financial services*</td>
<td>FS</td>
<td>13</td>
<td>Chemicals</td>
<td>CH</td>
</tr>
<tr>
<td>4</td>
<td>Real estate*</td>
<td>RE</td>
<td>14</td>
<td>Basic resources</td>
<td>BR</td>
</tr>
<tr>
<td>5</td>
<td>Construction &amp; materials</td>
<td>CM</td>
<td>15</td>
<td>Oil &amp; gas</td>
<td>OG</td>
</tr>
<tr>
<td>6</td>
<td>Industrial goods &amp; services</td>
<td>IGS</td>
<td>16</td>
<td>Telecommunication</td>
<td>TC</td>
</tr>
<tr>
<td>7</td>
<td>Automobiles &amp; parts</td>
<td>AP</td>
<td>17</td>
<td>Health care</td>
<td>HC</td>
</tr>
<tr>
<td>8</td>
<td>Food &amp; beverage</td>
<td>FB</td>
<td>18</td>
<td>Technology</td>
<td>TG</td>
</tr>
<tr>
<td>9</td>
<td>Personal &amp; household goods</td>
<td>PHG</td>
<td>19</td>
<td>Utilities</td>
<td>UT</td>
</tr>
<tr>
<td>10</td>
<td>Retail</td>
<td>RT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The financial sector variables.
FIGURE 3  Dynamics of total connectedness index and network BIC scores over the period 2007–2014, obtained from a rolling estimation with a window size of 100 days.

(a) Total connectedness index. (b) BIC of connectedness. The index of Granger causality is in blue and the Bayesian method is in red.

We compare the rank positions of the eigenvector centrality on the estimated networks. Eigenvector centrality is a measure of the importance of a variable in the transmission of systemic information and the spread of risks. (See Billio et al (2012) and Dungey et al (2012) for further discussions.) Figure 4 shows the evolution of the Spearman correlations over the sample period. The distribution of the correlations seems negatively skewed, with the extremely negative correlated rank recorded for the window ending November 12, 2008. Figure 5 shows the network of the window
FIGURE 4  Spearman correlations of centrality rank on the estimated networks over the sample period.

FIGURE 5  Volatility network for period ending November 12, 2008.

(a) GCnet (BIC = −3716.2). (b) BGnet (BIC = −4547.9). Edges are lagged dependencies.

ending November 12, 2008. The GCnet looks more connected than the BGnet; however, the associated BIC score favors the latter. Table 3 shows the top and bottom five ranked institutions from the two networks. We notice that utilities, chemicals and
TABLE 3  The top and bottom five super-sectors ranked by eigenvector centrality (EC) of the GCnet and BGnet for the period ending November 12, 2008.

<table>
<thead>
<tr>
<th>Rank</th>
<th>ID</th>
<th>Name</th>
<th>EC</th>
<th>ID</th>
<th>Name</th>
<th>EC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RT</td>
<td>Retail</td>
<td>0.3811</td>
<td>CH</td>
<td>Chemicals</td>
<td>0.4189</td>
</tr>
<tr>
<td>2</td>
<td>FB</td>
<td>Food &amp; beverage</td>
<td>0.3219</td>
<td>HC</td>
<td>Health care</td>
<td>0.3639</td>
</tr>
<tr>
<td>3</td>
<td>IN</td>
<td>Insurance</td>
<td>0.2940</td>
<td>UT</td>
<td>Utilities</td>
<td>0.2971</td>
</tr>
<tr>
<td>4</td>
<td>CM</td>
<td>Construction &amp; materials</td>
<td>0.2940</td>
<td>TC</td>
<td>Telecom</td>
<td>0.2821</td>
</tr>
<tr>
<td>5</td>
<td>MD</td>
<td>Media</td>
<td>0.2688</td>
<td>BR</td>
<td>Basic resources</td>
<td>0.2714</td>
</tr>
<tr>
<td>15</td>
<td>TG</td>
<td>Technology</td>
<td>0.1591</td>
<td>AP</td>
<td>Automobiles &amp; parts</td>
<td>0.1531</td>
</tr>
<tr>
<td>16</td>
<td>UT</td>
<td>Utilities</td>
<td>0.1591</td>
<td>RE</td>
<td>Real estate</td>
<td>0.1531</td>
</tr>
<tr>
<td>17</td>
<td>CH</td>
<td>Chemicals</td>
<td>0.0859</td>
<td>PHG</td>
<td>Personal &amp; household goods</td>
<td>0.0947</td>
</tr>
<tr>
<td>18</td>
<td>BR</td>
<td>Basic resources</td>
<td>0.0711</td>
<td>FB</td>
<td>Food &amp; beverage</td>
<td>0.0731</td>
</tr>
<tr>
<td>19</td>
<td>OG</td>
<td>Oil &amp; gas</td>
<td>0.0414</td>
<td>TL</td>
<td>Travel &amp; leisure</td>
<td>0.0731</td>
</tr>
</tbody>
</table>

Basic resources are ranked low in the GCnet but high in the BGnet. Also, food & beverage ranks highly in the GCnet but lowly in the BGnet.

Finally, we focus on the centrality of the financial sector of the market. Figures 6(a) and 6(b) show the dynamics of the rank differences and the most central institutions. Figures 6(c) and 6(d) also show the frequency of the sign rank differences and the frequency of the most central institution, respectively. We remind the reader that in rank terms, 1 means higher centrality and 4 means lower centrality. Also, the rank difference is the difference between the rank on the GCnet and that on the BGnet. Thus, a negative rank difference denotes a higher (respectively, lower) centrality on the GCnet (respectively, BGnet). The evolution of the rank differences in Figure 6(a) shows many deviations from the reference line (in blue) for all four institutions. The reference line indicates equal centrality ranks of institutions by the two estimations.

Figure 6(b) shows many periods of differences in the most central institution predicted by the two methods. For instance, GCnet found real estate to be more central than banks for most of 2007 to 2009, while the BGnet shows the opposite. The frequency of the sign rank differences (in Figure 6(c)) indicates results that are the complete opposite, such that, in most of the sample, banks and insurers are ranked relatively low on GCnet and high on the BGnet. Similarly, financial services and real estate are ranked relatively high on GCnet and low on the BGnet. Thus, a researcher using Granger causality will identify financial services to be most central in the spread

www.risk.net/journal  Journal of Network Theory in Finance
FIGURE 6  Centrality in the financial sector of European stock market. Rank value 1 (4) means highest (lowest) centrality. [Figure continues on next page.]

(a) Institutional rank differences. (b) Most central institution: (i) Granger causality; (ii) Bayesian method. A negative (positive) sign of rank difference means a higher (lower) centrality by the GCnet and a lower (higher) centrality by the BGnet.
of risk, and banks the least, while a researcher applying the Bayesian method will find banks to be the most central, and real estate the least.

In many real-world interactions, dependencies among random variables are more complex than pairwise. The Bayesian approach to network selection discussed in this paper is designed to handle joint estimation and large-scale multiple testing problems. Thus, it produces more suitable graphs than Granger causality to analyze complex interactions. From the BGnet, we find evidence that banks and insurers are more
central in the “fear connectedness” expressed by market participants in the financial sector of the euro area.

7 CONCLUSION

This paper presents a state-of-the-art review of the interface between statistics and econometrics in the inference and application of Bayesian graphical models. We specifically highlight connections and possible applications of network models in financial econometrics in the context of systemic risk. Using the monthly return indexes of Billio et al. (2012) for hedge funds, banks, brokers and insurers, we find evidence of a higher connectedness between 2001 and 2008 in the US financial system. We also find evidence that insurers play a central role in the vulnerability of the system, which amplified the global financial crisis. Further empirical study on the financial super-sectors of the European stock market reveals banks and insurers to be central figures in the “fear connectedness” (Diebold and Yilmaz 2014) expressed by market participants in the euro area. A comparison of the estimated networks shows that the Bayesian method produces dependence patterns that are more suitable than those produced by Granger causality to capture complex interdependencies.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper. The authors’ research is supported by funding from the European Union, Seventh Framework Programme FP7/2007-2013 under grant agreement SYRTO-SSH-2012-320270; the Italian Ministry of Education; University and Research (MIUR) PRIN 2010-11 grant MISURA; and the Air Force Office of Scientific Research (AFOSR) grant.

ACKNOWLEDGEMENTS

I would like to thank the editor, the anonymous referees, and Monica Billio, Roberto Casarin, Francesco Ravazzolo and Fabrizio Leisen for their comments on an earlier version of this paper.

REFERENCES


www.risk.net/journal Journal of Network Theory in Finance


