ISSN: 2281-1346

UNIVERSITÀ DI PAVIA
Department of Economics and Management

DEM Working Paper Series

A Bayesian Covariance Graph And Latent Position Model For Multivariate Financial Time Series

Daniel Felix Ahelegbey
(Università di Pavia)

Luis Carvalho
(Boston University)

Eric D. Kolaczyk
(Boston University)

 # 181 (02-20)

Via San Felice, 5
I-27100 Pavia

economiaweb.unipv.it

A Bayesian Covariance Graph And Latent Position Model For Multivariate Financial Time Series

Daniel Felix Ahelegbey[a,b,*], Luis Carvalho[b], Eric D. Kolaczyk[b]

[a] *Department of Economics and Management, University of Pavia, Italy*
[b] *Department of Mathematics and Statistics, Boston University, USA*

Abstract

Current understanding holds that financial contagion is driven mainly by system-wide interconnectedness of institutions. A distinction has been made between systematic and idiosyncratic channels of contagion, with shocks transmitted through the latter expected to be substantially more likely to lead to a crisis than through the former. Idiosyncratic connectivity is thought to be driven not simply by obviously shared characteristics among institutions, but more by the latent strategic position of firms in financial markets. We propose a Bayesian hierarchical model for multivariate financial time series that characterizes the interdependence in the idiosyncratic factors of a VAR model via a covariance graphical model whose structure is modeled through a latent position model. We develop an efficient algorithm that samples the network of the idiosyncratic factors and the latent positions underlying the network. We examine the dynamic volatility network and latent positions among 150 publicly listed institutions across the United States and Europe and how they contribute to systemic vulnerabilities and risk transmission.

Keywords: Bayesian inference, Covariance graph model, Idiosyncratic Contagion Channels, Latent Space Models, Systemic Risk, VAR

JEL: C11, C15, C51, C52, C55, G01

1. Introduction

Current understanding holds that the global financial crisis (henceforth GFC) did not occur as a result of a single event but a cluster of crises that rippled through the financial system (Bernanke, 2013; Tang et al., 2010). The report by the Financial Crisis Inquiry Commission (2011) reveals that the impact of the GFC was driven by distress in one area of the financial markets, the housing market, that led to failures in other areas by way of interconnections and vulnerabilities that bankers, government officials, and others had missed or dismissed. The vulnerabilities, according to Bernanke (2013), are the pre-existing structural weaknesses of the financial system that amplified the initial shocks. There is a myriad of existing studies to uncover these vulnerabilities using network models to identify channels of shock transmission among financial institutions and markets (Acemoglu et al., 2015; Arregui et al., 2013; Battiston et al., 2012; Billio et al., 2012; Diebold and Yilmaz, 2014; Elliott et al., 2014; Ladley, 2013; Moghadam and Viñals, 2010). The common lesson learnt from the recent crises and empirical research is the importance in understanding the structure

*Corresponding author

of financial networks, how the interconnections are generated, and the effect of the network on the stability or fragility of the system.

Financial institutions are often interconnected through bilateral obligations like direct deposits, lending, derivatives, etc. They are also interconnected through exposures to common risk or market factors. Network of financial institutions can therefore be characterized by different links due to different modes of connection (Langfield and Soramäki, 2016). Analytical works on modeling these interconnectedness can be classified into two main streams depending on how network links are constructed. On one end, links emerge from analyzing balance sheet data of direct bilateral transactions (Cont et al., 2013; Georg, 2013; Halaj and Kok, 2015; Minoiu and Reyes, 2013). On the other end, network links are estimated from the co-movement in market-based data on security prices (see Ahelegbey et al., 2016a; Barigozzi and Brownlees, 2019; Basu et al., 2016; Billio et al., 2012; Diebold and Yilmaz, 2014; Hautsch et al., 2015). There is relatively little empirical work on the former, largely because constructing networks on bilateral transactions requires detailed balance-sheet data and financial statements, which are generally hard to obtain and have low update frequencies (at best, quarterly). There is however a countless number of empirical works on market-based networks due to easy access to public data. A common limitation with the latter is the inability to explain how the estimated links are formed between institutions. Also, such network estimation methods are often unable to incorporate strategic behavior of institutions (Chan-Lau, 2017). This paper advances a statistical technique to analyze market-based networks that incorporates strategic behavior of institutions.

Market prices of financial securities usually come in the form of time series observations. The commonest model adopted to approximate the dynamic interactions among asset returns or volatilities is the vector autoregressive (VAR) representation. This class of model presents a convenient framework to capture the serial correlation in the return or volatility of financial assets, and has been extensively applied to analyze financial networks (see Ahelegbey et al., 2016a; Barigozzi and Hallin, 2017; Basu et al., 2016; Billio et al., 2012; Diebold and Yilmaz, 2014). It is, however, well known that market prices that are meant to signal the performance of institutions reflect both market (public) and firm-level information about the institutions (Roll, 1988). Thus, the result of networks estimated from these observations depends strongly on the relative amount of market (public) and firm-level information capitalized into the measurements. When applying VAR models, a distinction is made between systematic (market driven) and idiosyncratic (firm driven) interconnections. The former is usually analyzed via temporal dependence using Granger-causality techniques (Barigozzi and Brownlees, 2019; Basu et al., 2016; Granger, 2008; Kock and Callot, 2015). The idiosyncratic network, on the other hand, is analyzed via contemporaneous interactions of the VAR errors using (partial) correlation-based approaches (Barigozzi and Brownlees, 2019; Battiston et al., 2012; Giudici and Spelta, 2016). It has been shown that risk propagation via idiosyncratic networks can have a far more severe impact on the financial stability of the system with a higher likelihood of a crisis than through systematic networks (see Dungey and Gajurel, 2015; Tang et al., 2010).

In this study, we advance the idea that links in idiosyncratic market-based networks can be explained not simply by observed shared characteristics, but more by unobserved strategic behaviors (proxied by latent position of institutions). We formulate a three stage hierarchical model. The first stage approximates the dynamics in volatility of financial securities as a VAR model. We fit the model to extract estimates of the idiosyncratic components. The second stage estimates the adjacency matrix among the idiosyncratic components via a covariance

graphical model (CGM). The final stage then models the adjacency matrix as a logistic regression approximated by a latent positions model (LPM). We estimate our hierarchical model using a Bayesian scheme. Thus, we refer to our approach as the Bayesian covariance graph and latent positions model (BCGLPM).

The application of latent position in financial networks is motivated by the idea that "position matters", i.e, the link between institutions is driven by their relative strategic positions. For instance, it is well known in finance that to reduce risk, institutions and other investors must diversify their investments. However, when the diversification strategy becomes similar among many market players, micro-level diversification leads to macro-level concentration which exposes institutions to common risk factors thereby creating a fragile system for the spread of risk. Thus, the probability of a link between two institutions increases as their latent positions become more similar.

The contributions of this work are manifold. Firstly, we contribute to the construction of volatility networks from large samples of security prices using rolling-window VAR models. Our application is particularly related to the works of Barigozzi and Brownlees (2019); Diebold and Yilmaz (2014); Engle et al. (2012). More precisely, our work is in the spirit of Diebold and Yilmaz (2014), who estimated the dynamics of volatility connectedness via a VAR approximated model by using a rolling estimation which is able to track the instability in the financial system. The main difference of our application is that we do not only study the dynamics of the network but also the latent positions that explains the links.

Secondly, we contribute to the growing research on large covariance selection. The covariance graphical approach considered in this study is different from concentration graphical models, often referred to as covariance selection (Dempster, 1972). Covariance graphs are estimated by zeros in the covariance matrix while concentration graphs are generated by zeros in the precision matrix. More so, concentration matrices represent conditional independence as opposed to covariance matrices that capture marginal independence. Thus, concentration graphs are more sensitive to the inclusion/exclusion of relevant variables than covariance graphs, given that some variables are in/out of the model. Although there are several methods for sparse concentration graph selection, only a few address sparse covariance estimation (see Bien and Tibshirani, 2011; Khare et al., 2011; Rothman et al., 2009; Silva and Ghahramani, 2009; Wang, 2015). We contribute to the literature on large sparse covariance estimation by building on the Bayesian method of Wang (2015).

Thirdly, we contribute to the application of latent positions in financial networks. Our work relates to a strands of literature ranging from eigen-decomposition of network matrices (Hoff, 2008), to eigenmodel for longitudinal relational data (Hoff, 2015), dynamic latent distance models (Sarkar and Moore, 2005; Sewell and Chen, 2015), and dynamic latent eigenmodels (Durante and Dunson, 2014). Another related but different stream of literature is the application of factor models in finance (see Dungey et al., 2005; Dungey and Gajurel, 2015; Forbes and Rigobon, 2002; Fox and Dunson, 2015; Lopes and Carvalho, 2007; Nakajima and West, 2013). The difference, however, is that in standard factor models, the factors analyze the underlying drivers of the observed time series, while that of network models provide insights on the strategic positions that explain the link between nodes.

Latent position models are usually applied in social network analysis where the network is known and the task is to estimate the latent positions that explain the network links (see Hoff, 2008; Hoff et al., 2002). In our application, the network is unknown. Thus, incorporating the latent positions increases the number of model parameters. Using a Bayesian technique, we design an efficient scheme that infers jointly the network structure and the latent positions.

Our scheme combines the covariance structure learning algorithm of Wang (2015) and the simulation of latent positions by Hoff (2009). The highlight of our algorithm is that it samples the network and latent positions iteratively, using information about the latter to update the network priors.

We assess the stability of the network structure and latent positions of 150 publicly listed institutions across the United States and Europe, covering 2002–2014. We find evidence of high interconnectedness among institutions that began to manifest during the eve of the global financial crisis (i.e early-2008) and the European sovereign debt crisis (i.e early-2011). Furthermore, we find that the vulnerability in the system that manifested during the financial crises actually started from early-2007 when many firms experienced liquidity shocks following the fall in housing prices and the abrupt shutdown in sub-prime lending. An important contribution of our application is that, by tracking the latent position, we find evidence that periods of high vulnerability are also characterized by high spatial clustering, which can be attributed to an increased similarity in strategic positions of institutions.

The rest of the paper proceeds as follows. In Section 2, we present a hierarchical model and a discussion of the model inference scheme. In Section 3, we provide an illustration of BCGLPM on synthetic datasets and a comparison with alternative approaches. Section 4 presents the empirical financial application and results, and Section 5 concludes the paper.

2. Bayesian Covariance Graph and Latent Positions Model

2.1. Hierarchical Model Formulation

This section presents a brief overview of the model formulation which consists of three stages as shown in Figure 1. The red-dashed rectangle represents the first stage, the green is second stage, and the blue is the third stage. The first stage is a VAR model that ap-

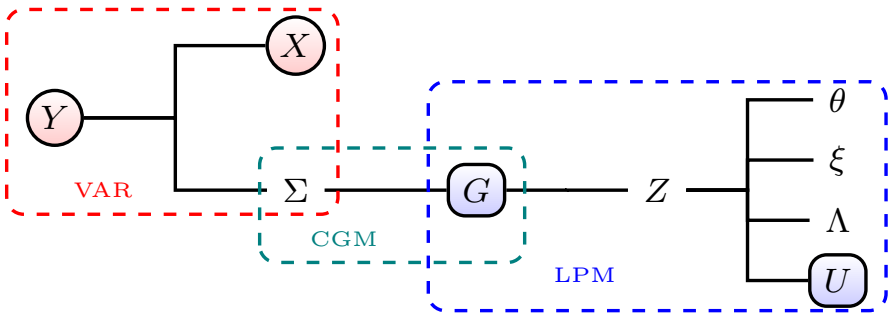

Figure 1: An illustration of the hierarchical model configuration. The red circled variables represent the observed data and the blue rectangle variables are our main parameters of interest.

proximates Y (a collection of the endogenous variables) as a function of X (a collection of past observations of the endogenous variables and common market indicators), and an idiosyncratic term whose covariance matrix is Σ. The second stage estimates an adjacency matrix of the network, G, associated with Σ via a covariance graphical model (CGM). The third stage models the adjacency matrix, G, as a logistic regression approximated by a latent positions model (LPM). This is done through a link a mapping function Z expressed in terms of $\{U, \Lambda, \theta, \xi\}$, obtained through eigen-decomposition of Z. More precisely, U is the latent coordinates (positions) matrix, Λ is an eigenvalue matrix, θ is a constant and ξ is an error term. Of all the parameters, our primary objective is inference on the idiosyncratic dependence structure, G, and the underlying latent coordinates, U.

2.1.1. VAR Model

Let $Y_t = (Y_{1,t}, \ldots, Y_{n,t})$ be an n-dimensional vector of endogenous variables at time t, $M_t = (M_{1,t}, \ldots, M_{m,t})$ is m-dimensional vector of common exogenous variables, and $E_t = (E_{1,t}, \ldots, E_{n,t})$, an n-dimensional vector of idiosyncratic factors. We consider the dynamics of Y_t as a stationary VAR(1) with exogenous covariates and error terms that follow a structural equation model. The approximating model is given by

$$Y_t = A_y Y_{t-1} + A_m M_{t-1} + E_t \ = AX_t + E_t, \qquad E_t \sim \mathcal{N}(0, \Sigma) \tag{1}$$

$$E_t = BE_t + \epsilon_t = (I_n - B)^{-1}\epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, Q) \tag{2}$$

where $A = (A_y, A_m)$ is an $n \times k$ matrix of coefficients, $k = n+m$, $X_t = (Y_{t-1}, M_{t-1})'$ is a $k \times 1$ vector of past observations, E_t is independent and identically normal with covariance matrix Σ, ϵ_t is a vector of structural errors independent and identically normal, with a diagonal covariance matrix Q, I_n is an identity matrix of order n, and B is an $n \times n$ coefficient matrix such that $B_{i,j}$ measures the contemporaneous effect of a shock of firm j on firm i. We assume the parameters $\{A_y, A_m, \Sigma, B, Q\}$ are time-independent within a fixed window. Thus, adopting a rolling-window estimation provides a dynamic inference for these parameters. The normality assumption of the VAR model stems from the fact that the logarithmic transformation of volatilities is approximately normal (see Andersen et al., 2003; Diebold and Yilmaz, 2014). From our experience with large VAR models, most coefficients of lags greater than one tend to be concentrated around zero. Despite this observation, the model can easily be extended to higher lag orders or different lag orders on the exogenous and endogenous variables.

2.1.2. Covariance Graph Model (CGM)

Let $Y = (Y_1', \ldots, Y_T')'$, $X = (X_1', \ldots, X_T')'$ and $E = (E_1', \ldots, E_T')'$ be a collection of Y_t, X_t and E_t over a fixed window of length T. We assume without loss of generality that E is matrix-normally distributed. Following the expression in (2), the covariance structure of E can be obtained as

$$\Sigma \ = \ (I - B)^{-1} Q (I - B)^{-1'} \tag{3}$$

where $D^{-1'}$ is the transpose of D^{-1}, $Q = \text{diag}(\sigma_{\epsilon_1}, \ldots, \sigma_{\epsilon_n})$, a diagonal matrix, and B is the contemporaneous coefficients matrix of the error model in (2). From standard maximum likelihood, given Σ, the elements in B and Q can be obtained by

$$B_{i,\pi_i} = \Sigma_{i,\pi_i} \Sigma^{-1}_{\pi_i,\pi_i} \qquad \text{and} \qquad Q_{ii} = \Sigma_{i,i} - \Sigma_{i,\pi_i} \Sigma^{-1}_{\pi_i,\pi_i} \Sigma'_{i,\pi_i} \tag{4}$$

where B_{i,π_i} is the vector of coefficients from a univariate linear regression of E_i on E_{π_i} (the predictors of E_i), Q_{ii} is the variance of ϵ_i, Σ_{i,π_i} is the covariance between E_i and E_{π_i}, and Σ_{π_i,π_i} is the covariance among E_{π_i}. Here E_i represents the i-th variable in E. The matrix B encodes the relationship between the reduced-form errors in the sense that $B_{ij} \neq 0$ if $E_j \to E_i$ and 0 otherwise. From the correspondence between B and Σ in (4), the marginal independence relationships between any pair E_i and E_j is such that,

$$E_i \leftrightarrow E_j \iff \Sigma_{i,j} \neq 0 \iff (B_{ij} \neq 0 \text{ and } B_{ji} \neq 0) \tag{5}$$

We define a Gaussian bi-directed graph model, G, on E which is determined by zeros in Σ (see Bien and Tibshirani, 2011; Khare et al., 2011; Rothman et al., 2009; Silva and Ghahramani,

2009). This establishes a relationship between Σ and G such that

$$\Sigma_{ij} = 0 \iff G_{ij} = 0, \qquad \text{and} \qquad \Sigma_{ij} \neq 0 \iff G_{ij} = 0 \tag{6}$$

2.1.3. Latent Position Model (LPM)

Let U be $n \times r$ coordinate matrix of n points in an r-dimensional system. We denote with $\boldsymbol{u}_i$, the i-th row of U (i.e., the i-th node position), and U_j as the $n \times 1$ vector containing the j-th column of U. Following Hoff (2008), we parameterize the ij-th entry of G via a probit link function, given by

$$G_{ij} = \mathbf{1}(Z_{ij} > 0) \tag{7}$$
$$Z_{ij} = \theta + \boldsymbol{u}_i' \Lambda \boldsymbol{u}_j + \xi_{ij} = \theta + (U \Lambda U')_{ij} + \xi_{ij} \tag{8}$$

where $\mathbf{1}(Z_{ij} > 0)$ is the indicator function, i.e., unity if $Z_{ij} > 0$ and zero otherwise, Z is the mapping function with $\theta \in \mathbb{R}$ as the constant, $(U \Lambda U')_{ij}$ is the i-th row and the j-th column of $(U \Lambda U')$, $\Lambda = \text{diag}\ (\lambda_1, \ldots, \lambda_r)$, is a diagonal matrix of eigenvalues, and ξ is a symmetric matrix of independent and identically distributed normal errors such that $\xi_{ij} = \xi_{ji}$.

As shown by Hoff (2008), the elements of Λ help identify the presence of homophily or stochastic equivalence—where nodes have similar relational patterns with other nodes in the network. Suppose the latent coordinates of node i are similar to those of j, i.e., $\boldsymbol{u}_i \approx \boldsymbol{u}_j$, and $U_{l,s} > 0,\ l = i, j$ and $s = 1, 2$, and that the effect of Λ is such that $\lambda_i > 0,\ i = 1, 2$. Then there is a tendency for nodes i and j to be connected. Thus, $\lambda_i > 0$ indicates homophily. Alternatively, if $\lambda_i < 0,\ i = 1, 2$, then there is a tendency for nodes i and j to be disconnected, although they share similar latent coordinates. Thus, $\lambda_i < 0$ indicates anti-homophily.

2.2. Parameters and Prior Specification

From the model discussed above, the parameters to estimate are $(A, \Sigma, G, Z, U, \Lambda, \theta, r)$. In considering large numbers of variables with relatively small sample size, we are confronted with the problem of over-parameterization. Estimating parameters jointly is a challenging problem and a computationally intensive exercise.

To handle the over-parameterization in large VAR models, various techniques discussed in the literature ranges from dynamic (sparse) factor and compression models (see Kaufmann and Schumacher, 2013; Koop et al., 2016; Lopes and Carvalho, 2007), to model selection and/or shrinkage methods (see Basu and Michailidis, 2015; George et al., 2008; Tibshirani, 1996), and graphical models for time series data (e.g., Ahelegbey et al., 2016a,b; Dahlhaus and Eichler, 2003; Dobra et al., 2004).

We follow the literature adopting a Bayesian paradigm that shrinks the coefficients in A while focusing on extracting estimates of the idiosyncratic component, E, and its associates covariance matrix Σ. This is made feasible by integrating out A with respect to its prior distribution to obtain a marginal likelihood function. Under the assumption that E follows a matrix-normal distribution, $E \sim \mathcal{MN}(\mathbf{0}, \Sigma, I_T)$, where $\mathbf{0}$ is a $T \times n$ zero matrix, Σ is the $n \times n$ row-specific covariance matrix, and I_T is the column-specific covariance matrix under the assumption that the idiosyncratic factors are independent over time. The matrix form of (1) can be expressed as $Y = XA' + E$. The conditional distribution of Y given A, Σ and X is $Y|A, \Sigma \sim \mathcal{MN}(XA', \Sigma, I_T)$, whose likelihood function is as follows:

$$P(Y|A, \Sigma) = (2\pi)^{-\frac{nT}{2}}\ |\Sigma|^{-\frac{T}{2}}\ \text{etr}\Big(-\frac{1}{2}[\Sigma^{-1}(Y - XA')'(Y - XA')]\Big) \tag{9}$$

where etr$(\cdot)$ is the exponential of the standard trace function.

2.2.1. Prior Distribution on A

We assume a matrix-normal conditional prior distribution of A given Σ, $A|\Sigma \sim \mathcal{MN}(A_0, \Sigma, \Psi)$ with the following density

$$P(A|\Sigma) = (2\pi)^{-\frac{1}{2}nk}\ |\Psi|^{-\frac{1}{2}n}\ |\Sigma|^{-\frac{1}{2}k}\ \text{etr}\Big(-\frac{1}{2}[\Sigma^{-1}(A-A_0)\Psi^{-1}(A-A_0)']\Big) \tag{10}$$

where A_0 is the prior mean of A, the row-specific prior covariance matrix of A is proportional to Σ, and the column-specific prior covariance matrix of A is proportional to Ψ. Note that A_0, Σ and Ψ are matrices of dimension $n \times k$, $n \times n$ and $k \times k$ respectively.

Proposition 1. *Under the prior $P(A|\Sigma)$ in (10), the likelihood $P(Y|A,\Sigma)$ in (9), and with $S_{xx} = X'X + \Psi^{-1}$, $S_{yx} = Y'X + A_0\Psi^{-1}$, $S_{yy} = Y'Y + A_0\Psi^{-1}A_0'$, and $S_{y|x} = S_{yy} - S_{yx}S_{xx}^{-1}S_{yx}'$, the marginal likelihood for any covariance matrix Σ is*

$$P(Y|\Sigma) \propto\ |S_{y|x}|^{-\frac{n}{2}}\ etr\big(-\frac{1}{2}\Sigma^{-1}S_{y|x}\big)\ . \tag{11}$$

Proof. Combining $P(Y|A,\Sigma)$ in (9) and $P(A|\Sigma)$ in (10), the marginalization of A with respect to its prior distribution is as follows:

$$\begin{aligned}
P(Y|\Sigma) &= \int_A P(Y|A,\Sigma)\ P(A|\Sigma)\ dA \\
&\propto \int_A \text{etr}\Big(-\frac{1}{2}\Sigma^{-1}\Big[A(X'X+\Psi^{-1})A' - 2(Y'X + A_0\Psi^{-1})A' + (Y'Y + A_0\Psi^{-1}A_0')\Big]\Big)dA \\
&\propto \int_A \text{etr}\Big(-\frac{1}{2}\Sigma^{-1}\Big[AS_{xx}A' - 2S_{yx}A' + S_{yy}\Big]\Big)\ dA \\
&\propto \int_A \text{etr}\Big(-\frac{1}{2}\Sigma^{-1}\Big[(A - S_{yx}S_{xx}^{-1})S_{xx}(A - S_{yx}S_{xx}^{-1})' + S_{yy} - S_{yx}S_{xx}^{-1}S_{yx}'\Big]\Big)dA.
\end{aligned} \tag{12}$$

From the expression above, the posterior distribution of A is matrix-normal:

$$A|Y,X,\Sigma \sim \mathcal{MN}(S_{yx}S_{xx}^{-1}, \Sigma, S_{xx}^{-1}) \tag{13}$$

Substituting $\hat{A} = S_{yx}S_{xx}^{-1}$ and $S_{y|x} = S_{yy} - S_{yx}S_{xx}^{-1}S_{yx}'$ into (12),

$$P(Y|\Sigma) \propto \text{etr}\Big(-\frac{1}{2}\Sigma^{-1}S_{y|x}\Big)\ \int_A \text{etr}\Big(-\frac{1}{2}\Sigma^{-1}(A-\hat{A})S_{xx}(A-\hat{A})'\Big)\ dA\ .$$

By definition $\int_A \text{etr}\Big(-\frac{1}{2}\Sigma^{-1}(A-\hat{A})S_{xx}(A-\hat{A})'\Big)\ dA\ = (2\pi)^{\frac{nk}{2}}|S_{xx}|^{-\frac{n}{2}}|\Sigma|^{\frac{k}{2}}$ and

$$P(Y|\Sigma) = (2\pi)^{-\frac{nT}{2}}\ |\Psi|^{-\frac{n}{2}}\ |S_{xx}|^{-\frac{n}{2}}\ |\Sigma|^{-\frac{T}{2}}\ \text{etr}\Big(-\frac{1}{2}\Sigma^{-1}S_{y|x}\Big)\ . \tag{14}$$

2.2.2. Prior Distribution on Σ and G

Following the spike and slab priors of Wang (2015), we assume an independent distribution on the off-diagonals and diagonals of Σ given by

$$\Sigma_{ij}|G\ \sim\ \mathcal{N}(0, V_{ij}),\quad \text{and}\quad \Sigma_{ii}|G\ \sim\ \text{Exp}\ \Big(V_{ii}/2\Big)$$

where V_{ij} is the ij-th column of V - an $n \times n$ symmetric hyper-parameter matrix of Σ. The above assumes that the off-diagonal elements of Σ are normally distributed, and the diagonals follows an exponential distribution. The density function is given by

$$P(\Sigma|G) \;= \prod_{i \neq j} \exp\left(-\frac{1}{2} V_{ij}^{-1} \Sigma_{i,j}^2 \right) \prod_{i=1}^{n} \exp\left(-\frac{1}{2} V_{ii} \Sigma_{i,i} \right) \mathbf{1}(\Sigma \in \mathcal{S}_+(G)) \tag{15}$$

where $\mathcal{S}_+(G)$ is the space of symmetric positive definite matrices with non-zero entries according to G. In this application, we define V_{ij} in a way that if $G_{ij} = 0$ then Σ_{ij} and the associated parameter $V_{ij} = v_0^2$ are concentrated around zero, and if $G_{ij} = 1$ then Σ_{ij} and $V_{ij} = v_1^2$ are different from zero. The choice of values for v_0 and v_1 is discussed in Section 2.3.

We consider the inclusion/exclusion of a link in G as a Bernoulli trial with density

$$P(G|U, \Lambda, \theta) \;= \prod_{i \neq j} \Gamma_{ij}^{G_{ij}} (1 - \Gamma_{ij})^{(1 - G_{ij})} \;, \tag{16}$$

$$\Gamma_{ij} \;=\; P(G_{ij} = 1 | U, \Lambda, \theta) \;=\; \Phi(\theta + (U \Lambda U')_{ij}), \tag{17}$$

where $\Gamma_{ij} \in (0, 1)$ is the probability of a link between nodes i and j, and Φ is the cumulative density function of the standard normal distribution.

2.2.3. Distribution on (θ, Λ, ξ, r, U, Z)

We specify prior distributions for the parameters θ, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_r)$, ξ, and U as:

$$\begin{aligned} \theta \;&\sim\; \mathcal{N}(\theta_0, \tau_\theta^2), & \lambda \;&\overset{iid}{\sim}\; \mathcal{N}(0, \tau_\lambda^2), & U &\sim \text{BMF}(C, D), \\ r \;&\sim\; \text{Unif}\,(\underline{r}, \bar{r}), & \xi_{ij} \;&\overset{iid}{\sim}\; \mathcal{N}(0, 1), & Z|G \;&\sim\; \mathcal{N}(\theta + U \Lambda U', \Sigma_Z) \end{aligned}$$

where $\theta_0, \tau_\theta, \tau_\lambda, \underline{r}, \bar{r}, C$, and D are the hyper-parameters. The above specification means that the parameters θ, ξ_{ij}, and $\lambda = (\lambda_1, \lambda_2)$ are normally distributed. Finally, we assume the columns of U are orthonormal and uniformly distributed in a r-dimensional space. The commonly used distribution for U is the Bingham-von Mises-Fisher (BMF) on the Stiefel manifold, $\mathcal{V}_{r,n} = \{U \in \mathcal{R}^{n \times r} : U'U = I\}$ (Bingham, 1974; Fisher, 1953; Khatri and Mardia, 1977). The associated probability density function of U is given by

$$P(U|C, D) \;\propto\; \text{etr}\Big(DU'CU\Big) \tag{18}$$

where C is $n \times n$ symmetric matrix and D is $r \times r$ diagonal matrix. The link function Z is normally distributed with mean $(\theta + U\Lambda U')$, and whose entries are constrained to be positive or negative depending on G. The density function of Z is given by:

$$P(Z|U, \Lambda, \theta) \;\propto\; \text{etr}\Big[-\frac{1}{4} (Z - \theta \mathbf{1}\mathbf{1}' - U\Lambda U')'(Z - \theta \mathbf{1}\mathbf{1}' - U\Lambda U') \Big] \tag{19}$$

where $\mathbf{1}$ is a vector of ones. The off-diagonals of Z have unit variances and the diagonals have variances of 2.

2.3. Setting Hyperparameters

For the hyperparameters governing the distribution of θ, i.e., $(\theta_0, \tau_\theta^2)$, we notice that θ_0 is positively related the link probabilities, Γ_{ij}. The expression of Z in (8) can be viewed as a

penalized bilinear similarity matrix function with θ as the penalty term. Thus, negative values of θ imply a lower value in Z and lower probability of the existence of links. We consider the problem of estimating a graphical model for large numbers of variables as one that can be approximated by a sparse structure. Thus, one would expect that the prior expectation of θ is negative. For this application, we set $\theta_0 = \Phi^{-1}(2/(n-1))$ and $\tau_\theta^2 = 100$. The choice of a θ_0 is a prior that is consistent with sparsity constraint in many application of Gaussian graphical models (see Wang, 2015).

For the spike and slab parameters, we follow standard applications by assuming v_0 to be small, concentrated around zero, and v_1 as a scaled version of v_0 and different from zero. By standardizing all datasets in our applications, we set $v_1 = 1$ and vary $v_0 = \{0.02, 0.05\}$.

Following Hoff (2009), we set $\tau_\lambda^2 = n$. As argued by the author, this value reflects the variance of the eigenvalues of a $n \times n$ matrix of independent standard normal noise.

We assume the prior expectation of A is a zero matrix, $A_0 = \mathbf{0}$ and the coefficients in A are a-priori independent within and across equations. Thus, $\Psi = \eta^{-1} I_k$ is a diagonal matrix, where η^{-1} is the prior variance and k is the number of covariates in X. Following our choice of $A_0 = \mathbf{0}$, the posterior expectation of A in (13) becomes $\hat{A}' = (X'X + \eta I_k)^{-1} X'Y$, which is the same as the ridge estimator with η as the ridge parameter. We set $\eta = c_0 k$, where c_0 varies on a grid between 0.1 to 10. To determine the optimal choice of c_0, we divide the data into 80% for estimation of $\hat{A}'$ and 20% for point forecast evaluation. In order not to have the posterior distribution of A dominated by its prior variance, we chose a grid, c_0, where the first difference of the mean squared forecast error (MSFE) is less than a tolerance level, e.g., 0.1. This is to avoid overfitting the data as well as not shrinking all coefficients to zero.

2.4. Choice of Dimension

For choice of the dimension of the latent space, we adapt the approach by Handcock et al. (2007) such that for each $r \in [\underline{r}, \bar{r}]$, we estimate a marginal likelihood of $P(Y, \hat{G}, \hat{U}|r)$ with $\hat{G}$ and $\hat{U}$ being the estimated network and latent position of the nodes. Thus, we approximate the BIC for the selection of r by

$$BIC = BIC_{\text{GM}} + BIC_{\text{PR}} + BIC_{\text{POS}} \tag{20}$$

where BIC_{GM} is the BIC of the graphical model, BIC_{PR} is BIC of the probit regression model, and BIC_{POS} is the BIC of the latent position. The BIC of the graph is given by

$$BIC_{\text{GM}} = -2 \log P(Y|\hat{\Sigma}, \hat{G}) + |\hat{G}| \log T \tag{21}$$

where $\hat{G}$ is the candidate graph, $|\hat{G}|$ is the number of links in the network, T is the number of observations, and $\log P(Y|\hat{\Sigma}, \hat{G})$ denotes the maximum likelihood function of the associated candidate graph model. We approximate the BIC of the probit regression model by

$$BIC_{\text{PR}} = -2 \log P(\hat{G}|\hat{U}, \hat{\Lambda}, \hat{\theta}) + d_{\text{PR}} \log N_{\text{PR}} \tag{22}$$

where $(\hat{\Lambda}, \hat{\theta})$ are the maximum likelihood estimator for (Λ, θ) given $U = \hat{U}$, d_{PR} is the number of elements in (Λ, θ) and $N_{\text{PR}} = n(n-1)/2$. The BIC of the latent position model is given by

$$BIC_{\text{POS}} = -2 \log P(\hat{U}|\hat{C}, \hat{D}) + d_{\text{POS}} \log n \tag{23}$$

where $(\hat{C}, \hat{D})$ are the maximum likelihood estimator for (C, D) given $U = \hat{U}$, d_{POS} is the number of elements in (C, D).

2.5. Bayesian Model Inference

Given the data, Y, and the prior distributions on the parameters, we are particularly interested in the posterior inference of the graph structure, G, and the underlying coordinates matrix, U. We proceed by performing the necessary posterior estimation through application of a Gibbs sampler that consists of the following steps:

$$[\Sigma|G, Y], \quad [G|\Sigma, U, \Lambda, \theta], \quad [Z|G, U, \Lambda, \theta], \quad [\theta|Z, U, \Lambda], \quad [\Lambda|Z, U, \theta], \quad \text{and} \quad [U|Z, \Lambda, \theta].$$

Note that the second Gibbs step is collapsed, that is, we marginalize Z. We sample $\{\Sigma, G\}$ following the results from Wang (2015) and $\{Z, \theta, \Lambda, U\}$ following Hoff (2009). A detailed description on how to sample the parameters is available in the online supplemental material.

3. Simulation Experiments

We evaluate the efficiency of our inferential approach on simulated datasets under the following scenarios:

$$\begin{aligned} &\textit{Lag-0 Setup}: \; Y_t = E_t, \quad E_t \sim \mathcal{N}(0, \Sigma_G) \\ &\textit{Lag-1 Setup}: \; Y_t = AY_{t-1} + E_t, \quad E_t \sim \mathcal{N}(0, \Sigma_G) \end{aligned}$$

where Σ_G is a covariance matrix constrained by a sparse graph G, such that $\Sigma_{ij} = 0$ if $G_{ij} = 0$ and $\Sigma_{ij} \neq 0$ if $G_{ij} = 1$. The *Lag-0 Setup* generates data from a regular Gaussian covariance graph model and the *Lag-1 Setup* incorporates an AR(1) temporal dependence on the endogenous variables with a Gaussian covariance graph structure on the residuals.

We generate edges in G from independent Bernoulli distributions with probability 0.2. We construct A to be a diagonal matrix. Following the random graph pattern of Wang and Li (2012), we construct $\Sigma_G = B_G + \delta I_n$, where B_G is a symmetric matrix constrained by G. We generate B_G, A and δ as follows:

$$(B_G)_{ij} = \begin{cases} 1 & \text{if } i = j \\ \beta_{ij} & \text{if } G_{ij} = 1 \\ 0 & \text{otherwise} \end{cases}, \quad A_{ij} = \begin{cases} a_i & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}, \quad \delta = \frac{n \min(\delta_b) - \max(\delta_b)}{1 - n}$$

where δ_b is the eigenvalues of B_G. The values of $(\beta_{ij} = \beta_{ji})$ and a_i are randomly drawn from a uniform distribution on $(-0.9, -0.3) \cup (0.3, 0.9)$. For the *Lag-1 Setup*, we initialize $Y_0 \sim \mathcal{N}(0, I_n)$. For each setup, we generate a dataset of dimension $n \in \{50, 100, 150\}$ and sample size $T = 2n$. We replicate the simulation and estimation exercise 10 times.

3.1. Competing Methods

Since there are no existing approaches for joint inference of the graph and latent positions, it is reasonable to compare our approach with closely related methods suitable for large covariance selection. To the best of our knowledge, the stochastic search structure learning (SSSL) method by Wang (2015) appears to be a suitable benchmark to compare our graph inference performance since it has been shown to be effective in dealing with large problems

and complicated models. Following the suggestion by the author of that paper, we set $v_0 = 0.02$ and the hyper-parameter for the graph priors, $\pi = 2/(n-1)$.

We present two versions of our BCGLPM approach for the inference of the graph. We denote by BCGLPM(0), the lag-0 version of the BCGLPM where the underlying approximated model is a structural equation model with a Gaussian covariance graph structure. Similarly, we denote by BCGLPM(1), the lag-1 version of the BCGLPM which assumes that the data is simulated from a VAR(1) model with idiosyncratic components that follow a Gaussian covariance graph structure. In both cases, our goal is to infer the graph structure and the latent nodal positions. To implement the BCGLPM(0) means setting X to be a null matrix which leads to replacing $S_{y|x} = S_{yy} = Y'Y$ in (14). Since the BCGLPM(0) is closely related to the SSSL, the former is expected to be very competitive against the latter.

3.2. Convergence and Mixing of MCMC

We run 10,000 MCMC iterations for all competing methods with the first 3,000 as the burn-in sample. All computations were implemented in MATLAB through the Boston University Shared Computing Cluster. We examine the mixing of the chains generated by the Gibbs sampler for the samples of θ, Λ, U, G, and Z. We monitor the mixing of the MCMC by computing the negative log likelihood score:

$$-2\log\mathcal{L}(Y|\mathcal{M}) \;=\; -2\log P(Y|\Sigma, G) \;-\; 2\log P(Z|\theta, U, \Lambda)\;. \tag{24}$$

We use the score to compute the potential scale reduction factor (PSRF) of Gelman and Rubin (1992). The chain is said to have converged if $PSRF \leq 1.2$. Details of the mixing of these quantities are provided in the supplementary material.

We estimate the posterior probability of the edges by averaging over the sampled networks, i.e., $\hat{\gamma}_{ij} = \frac{1}{H}\sum_{h=1}^{H} G_{ij}^{(h)}$, where H is the total number of iterations after the burn-in sample. Following the uncertainty in the determination of the presence/absence of a link between nodes in the network, we consider a one-sided posterior credibility interval for the edge posterior distribution. Using the criterion in Ahelegbey et al. (2016a), we parameterize the ij-th entry of the estimate of $\hat{G}$ via a link function:

$$\hat{G}_{ij} = \mathbf{1}(q_{ij} > c), \qquad\qquad q_{ij} = \hat{\gamma}_{ij} - z_{(1-\alpha)}\sqrt{\frac{\hat{\gamma}_{ij}(1-\hat{\gamma}_{ij})}{n_{\text{eff}}}} \tag{25}$$

where n_{eff} is the number of effective sample size, and $z_{(1-\alpha)}$ is the z-score of the normal distribution at $(1-\alpha)$ significance level. The default for c is 0.5, α is 0.05 and $z_{(1-\alpha)} = 1.65$.

3.3. Choice of Dimension

We report in Table 1 the BIC and average run time for every 100th iteration to choose the optimal dimension of the latent space for $r \in \{2, \ldots, 5\}$. The result is obtained by fixing $\theta_0 = \Phi^{-1}(2/(n-1))$, $\tau_\theta^2 = 100$, $\tau_\lambda^2 = n$, $v_0 = 0.02$ and varying $n \in \{50, 100, 150\}$ with $T = 2n$. For each different n, the optimal number of latent space dimension chosen by the BIC is $r = 2$. The result further shows that updating the positions of nodes in a two-dimensional latent space records the fastest run time for different simulation experiments. We, therefore, use $r = 2$ for our study which seems to agree with most applications involving multidimensional scaling and provides a convenient framework for the network visualization.

		$r=2$	$r=3$	$r=4$	$r=5$
$n=50$	BIC	**10464.59**	10601.32	10690.58	10799.78
	Time	34.20	38.85	44.71	48.03
$n=100$	BIC	**45061.72**	45265.89	45466.36	45671.08
	Time	145.62	161.10	176.76	191.75
$n=150$	BIC	**105292.19**	105575.64	105802.00	106154.28
	Time	1206.59	1283.11	1363.88	1430.30

Table 1: Comparing model selection for choice of r. Boldface values indicate the best choice for each metric.

3.4. Graph Predictive Evaluation

We analyze the network edge predictive performance of the competing methods in terms of graph accuracy (ACC) and the area under the receiver operator characteristic curve (AUC). The AUC depicts the true positive rate (TPR) against the false positive rate (FPR) depending on some threshold. TPR is the number of correct positive predictions divided by the total number of positives. FPR is the ratio of false positives predictions overall negatives. We report the average TP (true positives), FP (false positives), ACC and AUC for each of the competing methods.

3.5. Results

Table 2 presents the comparison of the graph performance of the simulation exercise. The result shows that higher values of v_0 produces more sparse networks with lower number of TP and FP. Furthermore, $v_0 = 0.02$ performs better at predicting higher number of true links depicted in the DGP than that of $v_0 = 0.05$. The overall predictive accuracy (i.e ACC and AUC) of the former is consistently higher than the latter.

In the following, we compare only BCGLPM(0) and BCGLPM(1) under $v_0 = 0.02$ with SSSL as the benchmark. When the true DGP is a lag-0 setup, that is without temporal dependence, we expect BCGLPM(0) and SSSL to outperform the BCGLPM(1). We, however, notice that overall, the estimated networks of the BCGLPM(0) and BCGLPM(1) appear more dense than the SSSL, with a higher average number of true predicted links and higher predictive accuracy. This is evidence that modeling network links via latent positions and the ability to update graph priors with latent position information produces better network structures than the benchmark method with fixed graph prior.

When the true DGP is a lag-1 setup, both SSSL and BCGLPM(0) are designed to estimate the network ignoring the temporal dependence in the underlying model. The result, therefore, shows that by ignoring the temporal dependence, the SSSL and BCGLPM(0) overestimates the average number of false predictions. Overall, the BCGLPM(1) infers a more accurate network structure and outperforms the SSSL and the BCGLPM(0) in both simulation setups. This provides evidence in favor of modeling idiosyncratic dependence in multivariate time series data using our proposed method.

4. Financial Connectedness and Latent Position Analysis

We analyze the interconnectedness and latent positions of top listed institutions in the financial sector of the United States and 15 European countries. We obtain daily price indexes on 150 institutions covering January 2002 to August 2014. The institutions reported include

		SSSL	BCGLPM(0)		BCGLPM(1)	
Model	v_0	0.02	0.02	0.05	0.02	0.05
Lag-0 Setup						
$n = 100$, $T = 2n$	TP	32.60	170.40	98.30	**187.50**	114.20
	FP	**1.00**	17.90	4.20	46.60	15.10
	ACC	80.64	**83.08**	81.90	82.84	82.00
	AUC	52.55	63.16	60.64	**63.59**	61.01
$n = 150$, $T = 2n$	TP	63.70	320.90	123.70	**373.30**	148.30
	FP	**0.90**	26.50	1.90	72.20	10.20
	ACC	80.66	82.73	81.18	**82.79**	81.33
	AUC	52.64	61.63	56.57	**62.29**	57.32
Lag-1 Setup						
$n = 100$, $T = 2n$	TP	30.30	**84.60**	60.70	80.00	49.10
	FP	26.00	88.10	49.80	21.60	**9.00**
	ACC	80.08	79.92	80.21	**81.17**	80.80
	AUC	51.81	54.72	54.25	**56.57**	55.40
$n = 150$, $T = 2n$	TP	51.20	**172.90**	89.90	154.10	60.80
	FP	41.00	177.00	73.20	33.60	**5.70**
	ACC	79.97	79.85	80.03	**80.96**	80.38
	AUC	51.38	54.26	53.06	**55.66**	53.37

Table 2: Model performance for different scenarios. Boldface values indicate the best choice for each metric.

50 banks (25 from US, 25 from Europe), 50 insurance companies (25–US, 25–Europe), 40 real estates companies (20–US, 20–Europe), 4 other top US banks that have been acquired or bankrupted (like Bear Sterns, Countrywide Financial Corporation, Lehman Brothers and Merrill Lynch) and 6 major global market indexes (S&P 500, Dow Jones, Nasdaq Composite, Euro Stoxx 600, Euro Stoxx 50 and HangSeng). Data on the global market indexes were obtained from Yahoo finance and the rest from Datastream. Countries represented by these institutions in Europe are Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, Greece, Ireland, Italy, the Netherlands, Norway, Sweden, and the UK. A detailed description of the data is available in the online supplemental material.

Let $P_{i,t}$ be the price index of institution i on day t, and $R_{i,t} = 100(\log P_{i,t-1} - \log P_{i,t})$ be the log-return. A preliminary analysis of the sample shows that out of the 150 institutions, 64% have serially correlated returns. Given the autoregressive behavior of the return series, we proxy the daily volatility of the prices by absolute first difference of the daily returns:

$$\sigma_{i,t} \approx \left| R_{i,t} - R_{i,t-1} \right| \tag{26}$$

This is based on the idea that, if returns at time $t-1$ is a prediction of the returns at time t, then $\sigma_{i,t}$ measures the absolute daily prediction error (deviation). For robustness check, we proxy $\sigma_{i,t}$ using a GARCH(1,1) model.

We evaluate the performance of our model, BCGLPM(1), against the BCGLPM(0) and the SSSL of (Wang, 2015) as a benchmark. Note that the BCGLPM(0) and the SSSL estimates the network of the log volatilities, Y_t, via a structural equation model setup ignoring the

serial correlated feature of volatilities of security prices. The difference, however, is that unlike the SSSL, the BCGLPM(0) adopts the latent position model to explain the network links and to update the graph priors. The BCGLPM(1), on the other hand, estimates a VAR approximated model via a Bayesian ridge-regularization and infers jointly the network of the idiosyncratic component and the latent positions underlying the network. To implement the competing methods, we set the prior hyperparameters as discussed in the simulation exercise.

We characterize the dynamics of the networks via a yearly (approximately 240 trading days) rolling window. Our choice of window size is intended to capture the annual (12-months) dependence among the institutions. We set the increments between successive rolling windows to a one-month period. The first window of our study is from January 2002 – December 2002, followed by February 2002 – January 2003, and the last from September 2013 – August 2014. In all, we have 141 rolling windows. First, we study the evolution of the network topology in terms of network density and clustering (transitivity). We then analyze the clustering behavior of firms by incorporating information from the latent position and the network structure. We further examine the Procrustes similarity of the latent positions to monitor the strategic behavior of institutions over time.

4.1. Network Density

Let G be an n-node graph without self-loop. We characterize (through numerical summaries) the time-varying nature of interconnections by monitoring the network density, $D(G) = E(G)/\binom{n}{2}$, i.e., the number of estimated links in the network divided by the total number of possible links. We present in Figure 2, the evolution of the network densities obtained from

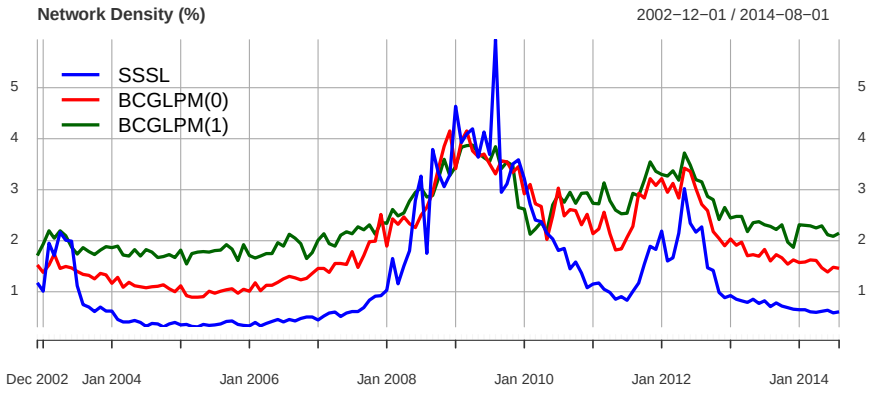

Figure 2: Network density of the competing methods via a 12-month rolling window estimation.

the yearly rolling window estimation. The BCGLPM(1) networks appear denser than the rest, while the BCGLPM(0) is moderately dense and that of the SSSL is very sparse. Denser networks have been shown to provide risk-sharing mechanism among firms as well as shock propagation (Acemoglu et al., 2015; Elliott et al., 2014; Glasserman and Young, 2016). The BCGLPM(1), thus, produces networks where shock propagations are more severe than the rest.

For comparison purposes, we define a Standardized Network Density Index (SNDI) — a measure of the degree of connectedness by rescaling the network density to a zero mean and a unit standard deviation. Figure 3 shows the SNDI for the competing methods. The zero-line of the plot can be viewed as the threshold line where negative values indicate bearable level

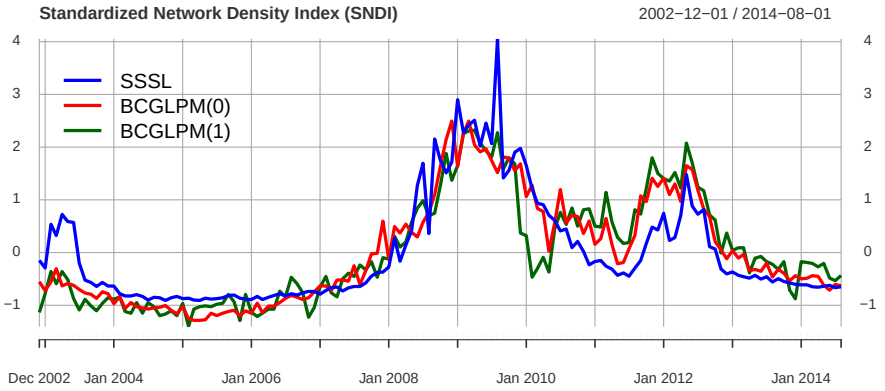

Figure 3: Standardized network density of the competing methods via a 12-month rolling window estimation.

of connectedness and positive values indicate a higher degree of system-wide vulnerability for risk propagation. From the result the periods of positive SNDI recorded by the SSSL are February 2008 – November 2010, and October 2011 – October 2012; that of BCGLPM(0) are December 2007 – April 2011, and July 2011 – January 2013; and that of BCGLPM(1) are February 2008 – January 2010, and June 2010 – March 2013. This shows that all the competing methods reveal that the crises times (global financial crisis (GFC) and the Euro area crisis) are characterized by high degree of system-wide vulnerabilities.

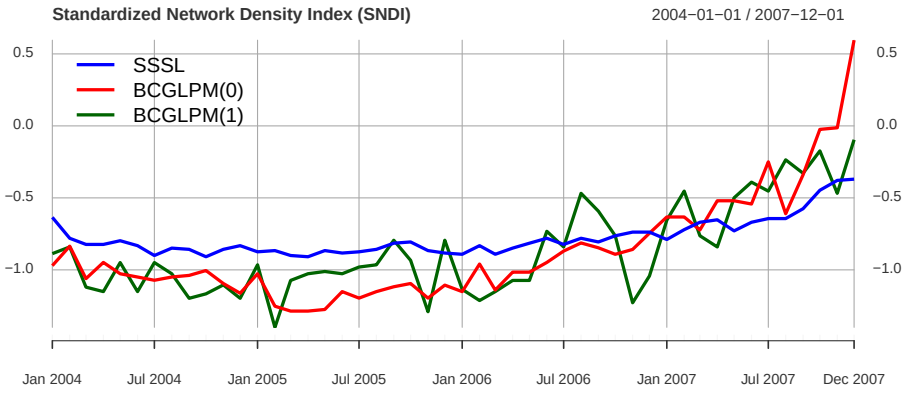

Figure 4: Standardized network density of the competing methods between January 2004 – December 2007.

We now focus on market events that preceded the GFC. In Figure 4, we depict the SNDI for the competing methods between 2004 to late-2007. We notice that all the methods recorded negative SNDI over the sub-period, which seem to indicate market "tranquillity" (or "calm before the storm"). A look the plots shows that, the SNDI of the SSSL is relatively stable over the sub-period. That of the BCGLPM(0) decreased during the early part of the period with a rise after 2006. The BCGLPM(1), on the other hand, shows a rise and fall in SNDI between 2004–2006 but a steady rise after 2006. This shows that, although the SNDI is negative over the sub-period, the calmness recorded by the SSSL can be misleading. Instead, we notice from the BCGLPM(0) and BCGLPM(1) that the vulnerability that began to manifest early-2008 did not just show up but was gradually mounting after 2006.

According to the timeline of market events, the years preceding the GFC saw a change

in the monetary policy of the Federal Reserve along with capital inflows to the US, creating conditions that allowed banks to lower rates and requirements. This gave way to easy lending even to consumers with less or no credit to acquire risky mortgages, in the anticipation that consumers would be able to quickly refinance at easier terms. The sub-prime mortgage origination between 2003 to 2006 increased from a low of 8.3% to an incredible level of 23.5%. Furthermore, the absence of regulatory actions contributed to the encouragement of many financial institutions holding highly related securities which led to a higher correlation among many investors and exposure to common risk factors. Within the sub-period under review, many top US institutions increased their leverages making them vulnerable to shocks. Around early- to mid-2007, many of these institutions experienced a liquidity shock following the fall in housing prices and abrupt shutdown of sub-prime lending. This led to losses for many financial institutions who held mortgage-related securities. Such events disrupted several market operations creating a cascade of the sale of securities by many institutions, lowering their values and increasing their volatility connections. These factors explain the steady rise in the vulnerabilities that began to manifest in the early period of 2008. More so, the near-collapse and acquisition of Bear Sterns by JP Morgan Chase in March 2008 further affected many institutions. The bankruptcy of Lehman Brothers (the fourth-largest U.S. investment bank at the time), and the bailout of American International Group (AIG - the world's largest insurance company) in September 2008, further triggered the actions of other market participants. This led to higher risk connections which amplified the shocks, affecting a broader aspect of the US financial system and many other correlated markets and economies.

4.2. Network Clustering Index

We study the dynamics of the degree to which financial institutions tend to cluster together and its relevance to the financial contagion process. We achieve this by monitoring the global clustering index of the estimated networks over the rolling windows. The clustering considered for this analysis follows Barrat and Weigt (2000) which corresponds to the social network concept of transitivity captured numerically as:

$$Cl(G) = \frac{3 \times (\text{number of triangles})}{(\text{number of open triads})} \tag{27}$$

where open triads are connected sub-graph consisting of three nodes and two edges. The index takes values between 0 and 1 and measures the tendency for nodes in a network form triangles. We observe from Figure 5 that the transitivity coefficient is quite high in turbulent times, with the index significantly higher than 0.3 in almost all cases.

4.3. Weighted Clustering Index

We extend the notion of average transitivity index to analyze the dynamics of institutions clustering behavior taking into account the latent positions. To fully characterize the interactions among the institutions, we considered representing the network of nodes as interconnections among spatial units by assigning weights to the edges using the cosine distance of the latent coordinates. Let $W = \{w_{ij}\}$ be $n \times n$ weighted adjacency matrix with $w_{ij} \in (0, 1)$ such that $w_{ij} \neq 0$ if $G_{ij} = 1$ and zero otherwise. The weighted clustering coefficient following

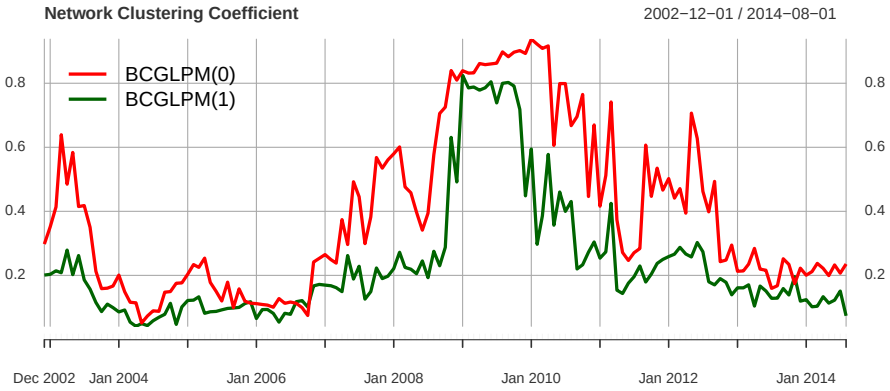

Figure 5: Network clustering coefficient of the BCGLPM(0) and BCGLPM(1), January 2002 – August 2014.

Barrat et al. (2004) is computed numerically by:

$$WC(G) \;=\; \frac{1}{n}\sum_{i} WC(i) \;=\; \frac{1}{n}\sum_{i}\left[\frac{1}{d_i(n-1)}\sum_{j,k}\frac{w_{ij}+w_{ik}}{2}G_{ij}G_{ik}G_{jk}\right]$$

where $WC(i)$ is the weighted clustering coefficient of node i, and $d_i = \sum_j w_{ij}$ is the weighted degree of node i. Figure 6 presents the time series of the weighted (average) clustering

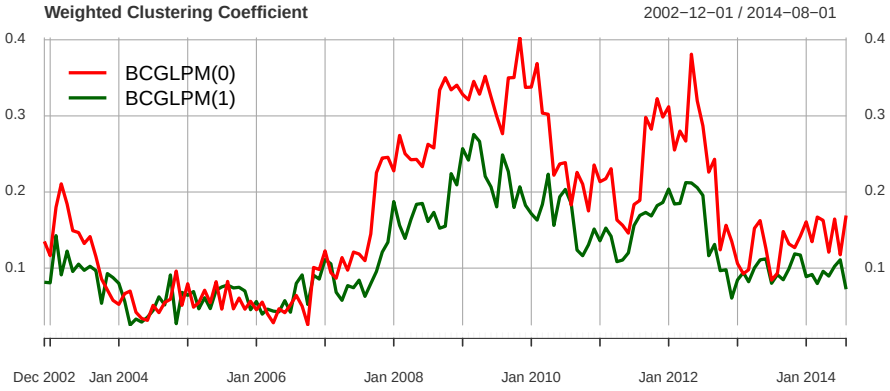

Figure 6: Weighted clustering coefficient of the BCGLPM(0) and BCGLPM(1), January 2002 – August 2014.

coefficients over the sample period. Though the range of the plot seems different from that of the unweighted global clustering coefficient in Figure 5, the dynamics are not significantly different. Thus, our result suggests that clustering among financial institutions becomes stronger in turbulent financial times and especially during the GFC and the Eurozone crisis period.

4.4. Procrustes Analysis

To investigate the strategic position of firms in financial markets, we adopted a Procrustean approach to analyze the latent positions underlying the estimated networks. It is well known that two different coordinates could produce positions that are very similar but look different due to rotation or scaling. The Procrustes function, therefore, transforms one

set of coordinates to make them comparable to the other through translation, uniform scaling, rotation and reflection (see Dryden and Mardia, 2016).

Let U^0 and U^1 denote the coordinates of the positions estimated by the BCGLPM(0) and BCGLPM(1) respectively. We compute $\hat{U}$ as the Procrustean transformation of U^1 with U^0 as the target. The transformation is generally expressed as:

$$\hat{U} = \rho U^1 H + c \tag{28}$$

where ρ is a scalar dilation, H is a 2×2 orthogonal matrix representing a rotation and reflection, and c is a 2×1 translation vector. In this application, we evaluate the similarity between two different coordinates of the nodal positions using Procrustes similarity metric, denoted by $S(U^0, U^1) = 1 - D(U^0, U^1)$ where $D(U^0, U^1)$ is the standardized distance between the target, U^0 and the transformed (i.e, $\hat{U}$). Like any similarity measure, values close to zero (one) indicate high dissimilarity (similarity) in the set of coordinates. Figure 7 shows the

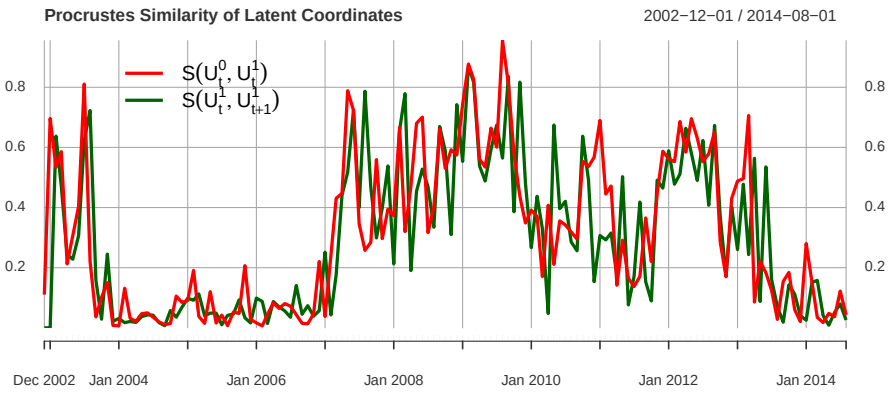

Figure 7: Similarity of the latent positions, $S(U_t^0, U_t^1)$ estimated by BCGLPM(0) and BCGLPM(1), and $S(U_t^1, U_{t+1}^1)$ measures the change in latent positions of BCGLPM(1) between two time periods.

dynamics of the similarity in the latent strategic position of financial institutions over the sample period. The result records the lowest levels of similarity in latent positions between late-2003 to mid-2007 and late-2013 to 2014, and the highest values between late-2008 to late-2009. Thus, the strategic position of financial institutions differ significantly during normal (or tranquil) periods and turbulent times.

We notice that the identified period of exceptionally low (high) Procrustes similarity seems to coincide with periods negative (positive)-SNDI. As discussed earlier, the losses in US sub-prime mortgages around early- to mid-2007 affected not just US institutions but other multinational financial institutions (e.g BNP Paribas - the largest French bank). This spurred up a change in strategic operations of other key market players after the signal from BNP Paribas in August 2007. Thus, many market players including investors and financial institutions stopped trading certain commodities, lenders quickly withdrew from many programs, and investors dumped their mortgage-backed security holdings and increased their holdings in seemingly safer money market funds and treasury bonds. Therefore in an attempt to reduce individual risk, many market players who pursue similar diversification and investment strategy end up creating a fragile system, vulnerable to the propagation of shocks. This to a large extent explains the rise in the similarity measure between late-2007 to late-2009 (the GFC period).

Further comparison with events in the global financial market shows that the rise in the similarity of the latent positions between 2010–2013 coincides with a time of struggle among Euro area members to recover from the global recession. As a result, the EU was thrown into a crisis that centered on heavily indebted countries like Greece, Ireland, Portugal, Spain, and Italy. We, therefore, notice a quick rise in the similarity among the institutions early-2010 and another from mid-2011 to late-2013. The latter sub-period can be explained by the threat to the EU financial institutions and the global financial system when the crisis of Greece, Ireland, and Portugal began to affect Spain and Italy (the third largest Euro area economy and second-biggest debtor to bond investors). More importantly, many European institutions were heavily exposed to Spain and Italy, thus, a sovereign default by these two countries will spread the crisis within and beyond Europe. This event spurred calls for changes in strategic operations among investors, sovereign bondholders and financial institutions.

4.5. Correlation of Network Statistical Measures

We turn our attention to the correlation among the different network statistical measures considered so far (see Table 3). SNDI(1), GClust(1) and WClust(1) in the table denote the standardized network density index, the global clustering index and the weighted clustering coefficient of the BCGLPM(1) networks respectively. The table shows that despite the differences in the various network statistical measures, the results are strongly positively correlated. More importantly, the similarity metric is positively related with network density and clustering. It is therefore safe to conclude that, periods of dense financial interconnections are characterized by high clustering among institutions which are attributable to similarities in strategic operations among counterparties.

	SNDI(1)	GClust(1)	WClust(1)	$S(U_t^1, U_{t+1}^1)$
SNDI(1)	1	0.8143	0.8967	0.6856
GClust(1)	0.8143	1	0.8227	0.5931
WClust(1)	0.8967	0.8227	1	0.6018
$S(U_t^1, U_{t+1}^1)$	0.6856	0.5931	0.6018	1

Table 3: Correlation of the Network Statistical Measures.

4.6. Network Visualization

To visualize the estimated network structure and latent node positions capture by our proposed scheme, we report in Figures 8 and 9 some selected yearly rolling windows. The figure shows both individual and group level representation of institutions. Selected institutions are represented by the abbreviation of their names. At group level, the nodes are color coded with shapes to classify the institutions into regional and institutional sectors. The triangle-shaped nodes represent European institutions and circle-shaped nodes for their US counterparts. Green-color nodes represent Banks, red for Insurance companies, blue for Real Estates, and cyan for currently non-existing institutions.

Unlike standard network graph representation, the layouts depicted in the figures are fully controlled by the latent coordinates jointly estimated with the network structure via the BCGLPM approach. The figure shows significant changes in the network configuration and node positions over the selected networks. We see evidence of spatial clustering among some regional and institutional sectors for sub-periods. For instance, the structure for the yearly rolling window ending December 2007 shows a group of European institutions, especially

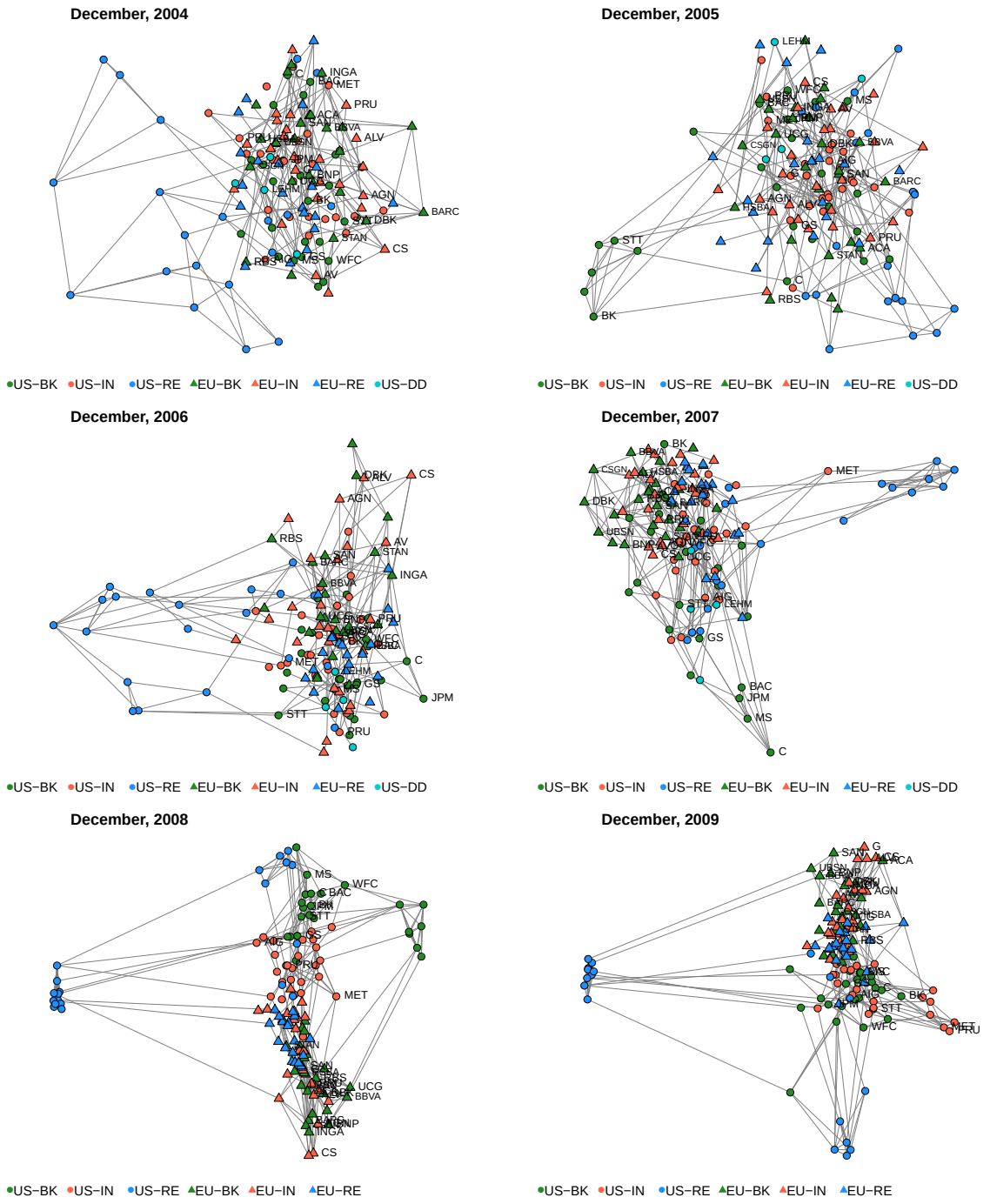

Figure 8: A representation of the network structure for the yearly rolling windows 2004–2009. US-BK is US Banks, US-IN for US Insurance Companies, US-RE for US Real Estates, EU-BK for EU Banks, EU-IN for EU Insurance Companies, EU-RE for EU Real Estates, and US-DD for US institutions currently not in existence.

banks (in green triangles) concentrated at the top left corner, and their US counterparts (in green circles) at the bottom, while US real estate companies (in blue circles) are gathered at the top right corner.

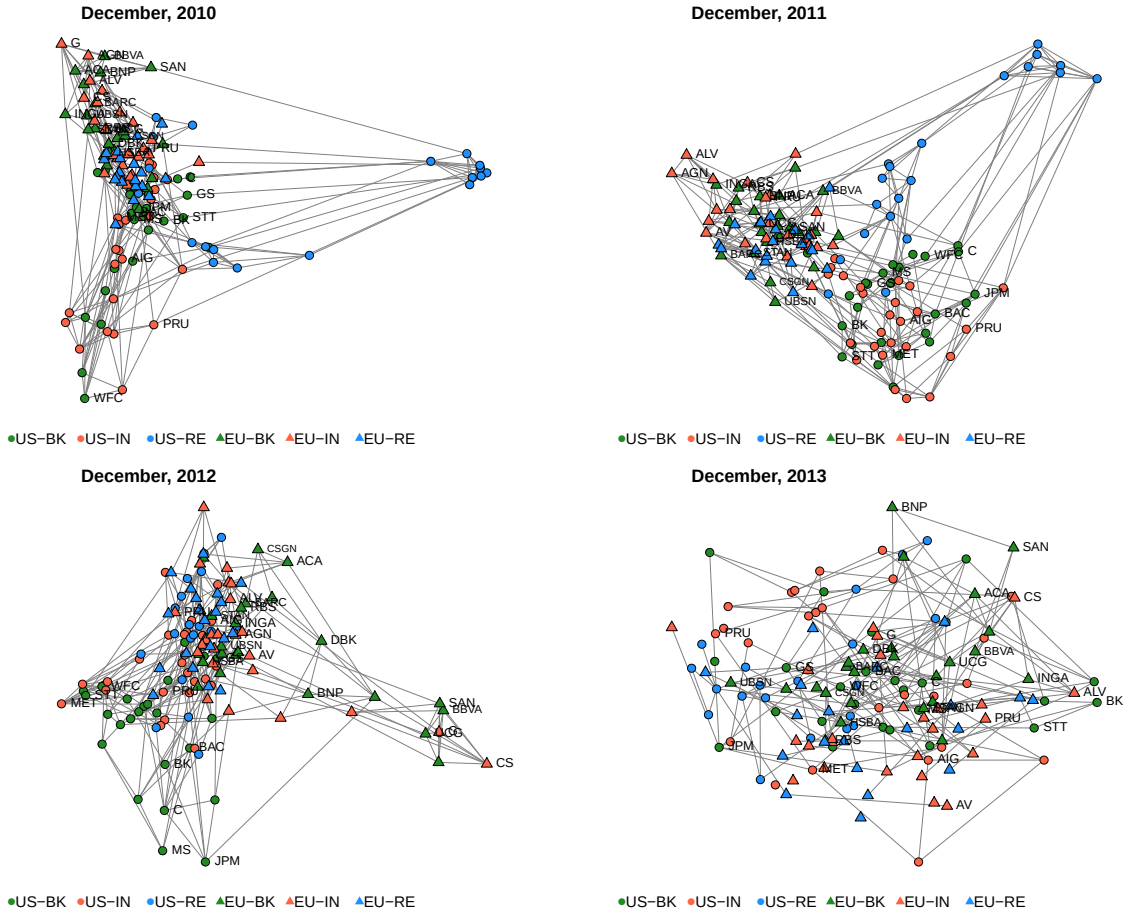

Figure 9: Network for years 2010–2013. US-BK is US Banks, US-IN for US Insurance Companies, US-RE for US Real Estates, EU-BK for EU Banks, EU-IN for EU Insurance Companies, and EU-RE for EU Real Estates.

The December 2008–2011 figures also depict a relatively different layout and a quite interesting network structure. From these graphs, we notice many US real estate institutions completely separated from the rest in terms of positioning though connected to themselves and other institutions. Thus, we detect strategic clustering behavior among institutions belonging to troubled sectors/markets during and after the financial crisis, providing evidence in favor of the use of latent position analysis in the estimation of financial networks.

5. Conclusion

In this paper, we study the strategic position of institutions in the construction of financial network from market-based data. We postulate that connectivity among institutions emerges not simply from the co-movement in observed security prices, but more from the similarity of the strategic behaviors proxied by latent factors. We, therefore, propose a hierarchical model that characterizes the idiosyncratic linkages among financial institutions by approximating the underlying multivariate volatility interactions as a VAR model. We estimate the approximated model via a Bayesian ridge-regularization and inferred jointly the network of the idiosyncratic dependence and latent positions underlying the network. We present an efficient Markov

chain Monte Carlo algorithm that samples the network and latent positions iteratively, using information from the latter to update the graph prior.

We demonstrate the effectiveness of our approach through simulation and empirical applications. We study the idiosyncratic volatility networks among 150 publicly listed institutions. The list of financial institutions is composed of banks, insurance companies, and real estates operating in the U.S. and Europe. The sample period covers January 2002 to August 2014. We proxy the daily volatility measure by the absolute first-order difference in daily log returns and conduct a robustness check using volatilities proxied by a GARCH process. Since there are no existing approaches for joint inference of the graph and latent positions, we compare the performance of our approach with the stochastic search structure learning method by Wang (2015), which appears to be a suitable benchmark due to its effectiveness in dealing with large covariance graph inference problems. We study the dynamics of the network and latent positions using a yearly rolling window estimation. A sensitivity analysis is conducted by considering a 2-year rolling window.

The result of the empirical analysis reveals that a higher level of interconnectedness and vulnerability of the system began to manifest early-2008. In analyzing the years preceding the global financial crisis, the result shows that 2004–2007 sub-period may not be as calm as many studies have suggested. Rather, by tracking and incorporating the latent position of institutions in the construction of the idiosyncratic networks, we find evidence that the vulnerability that began to manifest early-2008 did not just begin to show up on the eve of the crisis but actually started after 2006 when many firms started experiencing liquidity shocks following the fall in housing prices and the abrupt shutdown in sub-prime lending. Furthermore, we find that periods of high financial system vulnerability (proxied by high network densities) is also characterized by high clustering of institutions which can be attributed to a high level of similarities in strategic operations among counterparties. For instance, the events of 2008, beginning with the near-collapse of Bear Sterns, the bankruptcy of Lehman Brothers and the bailout of AIG triggered several reactions and disruptions in financial market activities. In an attempt to reduce individual risk, many market players pursued similar diversification and investment strategy which ended up creating a fragile system, vulnerable to the propagation of shocks. This led to higher clustering among some institutions belonging to some sectors with a rise in the similarities of their strategic operations. Also, the events of sovereign defaults in the Euro area during the sovereign crisis also triggered similar behavior among financial institutions. We demonstrate through graphical representation that incorporating the latent positions underlying the network reveals a strategic clustering behavior of firms such that institutions belonging to troubled sectors/markets appear separated from the rest (e.g real estates). This provided evidence in favor of the use of stochastic network and latent position models in analyzing financial contagion. As a recommendation to ensure a more robust and stable system, regulators must monitor both micro-level strategic behavior of institutions as well as macro-level concentration to avoid many key financial market players and investors becoming highly interconnected via exposure to common risk factors.

SUPPLEMENTARY MATERIAL

We present a detailed description of the sampling approach of the parameters and provide results of the mixing of the Markov chains generated by the Gibbs sampler for the samples of θ, Λ, the network structure, G, and Z in our simulation experiment. We conduct several

robustness checks to validate the sensitivity of our empirical results using: 1) a different measure of volatility as proxied by GARCH(1,1) of daily stock returns; 2) a different specification of the spike-and-slab parameter, i.e. $v_0 = 0.05$; 3) a two year rolling window for the model estimation; and 4) a different specification of the latent space dimension, ($r = 3$).

References

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015). Systemic Risk and Stability in Financial Networks. *American Economic Review 105*(2), 564–608.

Ahelegbey, D. F., M. Billio, and R. Casarin (2016a). Bayesian Graphical Models for Structural Vector Autoregressive Processes. *Journal of Applied Econometrics 31*(2), 357–386.

Ahelegbey, D. F., M. Billio, and R. Casarin (2016b). Sparse Graphical Vector Autoregression: A Bayesian Approach. *Annals of Economics and Statistics 123/124*, 333–361.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and Forecasting Realized Volatility. *Econometrica 71*(2), 579–625.

Arregui, N., M. Norat, A. Pancorbo, J. G. Scarlata, E. Holttinen, F. Melo, J. Surti, C. Wilson, R. Wehrhahn, and M. Yanase (2013). *Addressing Interconnectedness: Concepts and Prudential Tools.* International Monetary Fund.

Barigozzi, M. and C. Brownlees (2019). NETS: Network Estimation for Time Series. *Journal of Applied Econometrics 34*(3), 347–364.

Barigozzi, M. and M. Hallin (2017). A Network Analysis of the Volatility of High-Dimensional Financial Series. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 66*(3), 581–605.

Barrat, A., M. Barthelemy, R. Pastor-Satorras, and A. Vespignani (2004). The Architecture of Complex Weighted Networks. *Proceedings of the National Academy of Sciences 101*(11), 3747–3752.

Barrat, A. and M. Weigt (2000). On The Properties of Small-World Network Models. *The European Physical Journal B-Condensed Matter and Complex Systems 13*(3), 547–560.

Basu, S., S. Das, G. Michailidis, and A. Purnanandam (2016). Measuring Systemic Risk with Network Connectivity. In *Proceedings of the Second International Workshop on Data Science for Macro-Modeling*, pp. 11. ACM.

Basu, S. and G. Michailidis (2015). Regularized Estimation in Sparse High-dimensional Time Series Models. *The Annals of Statistics 43*(4), 1535–1567.

Battiston, S., D. D. Gatti, M. Gallegati, B. Greenwald, and J. E. Stiglitz (2012). Liaisons Dangereuses: Increasing Connectivity, Risk Sharing, and Systemic Risk. *Journal of Economic Dynamics and Control 36*(8), 1121–1141.

Bernanke, B. (2013). Monitoring the Financial System. Speech, at the 49th Annual Conference Bank Structure and Competition sponsored by the Federal Reserve Bank of Chicago, May 10.

Bien, J. and R. J. Tibshirani (2011). Sparse Estimation of a Covariance Matrix. *Biometrika 98*(4), 807–820.

Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2012). Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors. *Journal of Financial Economics 104*(3), 535 – 559.

Bingham, C. (1974). An Antipodally Symmetric Distribution on the Sphere. *The Annals of Statistics 2*(6), 1201–1225.

Chan-Lau, J. A. (2017). *Variance Decomposition Networks: Potential Pitfalls and a Simple Solution.* International Monetary Fund.

Cont, R., E. Santos, and A. Moussa (2013). Network Structure and Systemic Risk in Banking Systems. In *Handbook of Systemic Risk*, pp. 327–368. Cambridge University Press.

Dahlhaus, R. and M. Eichler (2003). Causality and Graphical Models for Time Series. In *Green, N. Hjort, and S. Richardson (eds.): Highly Structured Stochastic Systems.* University Press, Oxford.

Dempster, A. P. (1972). Covariance Selection. *Biometrics*, 157–175.

Diebold, F. and K. Yilmaz (2014). On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms. *Journal of Econometrics 182*(1), 119–134.

Dobra, A., C. Hans, B. Jones, J. R. Nevins, G. Yao, and M. West (2004). Sparse Graphical Models for Exploring Gene Expression Data. *Journal of Multivariate Analysis 90*(1), 196–212.

Dryden, I. L. and K. V. Mardia (2016). *Statistical Shape Analysis: With Applications in R.* John Wiley & Sons.

Dungey, M., R. Fry, B. González-Hermosillo, and V. L. Martin (2005). Empirical Modelling of Contagion: A Review of Methodologies. *Quantitative Finance 5*(1), 9–24.

Dungey, M. and D. Gajurel (2015). Contagion and Banking Crisis–International Evidence for 2007–2009. *Journal of Banking and Finance 60*, 271–283.

Durante, D. and D. B. Dunson (2014). Nonparametric Bayes Dynamic Modelling of Relational Data. *Biometrika 101*(4), 883–898.

Elliott, M., B. Golub, and M. O. Jackson (2014). Financial Networks and Contagion. *American Economic Review 104*(10), 3115–3153.

Engle, R. F., G. M. Gallo, and M. Velucchi (2012). Volatility Spillovers in East Asian Financial Markets: a MEM-based Approach. *Review of Economics and Statistics 94*(1), 222–223.

Financial Crisis Inquiry Commission (2011). *The Financial Crisis Inquiry Report: The Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States*. Public Affairs.

Fisher, R. A. (1953). Dispersion on a Sphere. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 217*(1130), 295–305.

Forbes, K. J. and R. Rigobon (2002). No Contagion, Only Interdependence: Measuring Stock Market Comovements. *The journal of Finance 57*(5), 2223–2261.

Fox, E. B. and D. B. Dunson (2015). Bayesian Nonparametric Covariance Regression. *The Journal of Machine Learning Research 16*(1), 2501–2542.

Gelman, A. and D. B. Rubin (1992). Inference from Iterative Simulation Using Multiple Sequences, (with discussion). *Statistical Science 7*, 457–511.

Georg, C.-P. (2013). The Effect of the Interbank Network Structure on Contagion and Common Shocks. *Journal of Banking and Finance 37*(7), 2216–2228.

George, E. I., D. Sun, and S. Ni (2008). Bayesian Stochastic Search for VAR Model Restrictions. *Journal of Econometrics 142*, 553–580.

Giudici, P. and A. Spelta (2016). Graphical Network Models for International Financial Flows. *Journal of Business & Economic Statistics 34*(1), 128–138.

Glasserman, P. and H. P. Young (2016). Contagion in Financial Networks. *Journal of Economic Literature 54*(3), 779–831.

Granger, C. W. (2008). Non-Linear Models: Where Do We Go Next - Time Varying Parameter Models? *Studies in Nonlinear Dynamics & Econometrics 12*(3).

Halaj, G. and C. Kok (2015). Modelling The Emergence Of The Interbank Networks. *Quantitative Finance 15*(4), 653–671.

Handcock, M. S., A. E. Raftery, and J. M. Tantrum (2007). Model-Based Clustering for Social Networks. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 170*(2), 301–354.

Hautsch, N., J. Schaumburg, and M. Schienle (2015). Financial Network Systemic Risk Contributions. *Review of Finance 19*(2), 685–738.

Hoff, P. D. (2008). Modeling Homophily and Stochastic Equivalence in Symmetric Relational Data. In *Advances in Neural Information Processing Systems*, pp. 657–664.

Hoff, P. D. (2009). Simulation of the Matrix Bingham-von Mises-Fisher Distribution, with Applications to Multivariate and Relational Data. *Journal of Computational and Graphical Statistics 18*(2), 438–456.

Hoff, P. D. (2015). Dyadic Data Analysis with AMEN. Technical Report 638, Department of Statistics, University of Washington.

Hoff, P. D., A. E. Raftery, and M. S. Handcock (2002). Latent Space Approaches to Social Network Analysis. *Journal of the American Statistical Association 97*(460), 1090–1098.

Kaufmann, S. and C. Schumacher (2013). Bayesian Estimation of Sparse Dynamic Factor Models with Order-Independent Identification. Working paper, Studienzentrum Gerzensee.

Khare, K., B. Rajaratnam, et al. (2011). Wishart Distributions for Decomposable Covariance Graph Models. *The Annals of Statistics 39*(1), 514–555.

Khatri, C. G. and K. V. Mardia (1977). The Von Mises-Fisher Matrix Distribution in Orientation Statistics. *Journal of the Royal Statistical Society. Series B (Methodological) 39*(1), 95–106.

Kock, A. B. and L. Callot (2015). Oracle Inequalities for High Dimensional Vector Autoregressions. *Journal of Econometrics 186*(2), 325–344.

Koop, G., D. Korobilis, and D. Pettenuzzo (2016). Bayesian Compressed Vector Autoregressions. Working paper, Social Science Research Network.

Ladley, D. (2013). Contagion and Risk Sharing on the Inter-bank Market. *Journal of Economic Dynamics and Control 37*(7), 1384–1400.

Langfield, S. and K. Soramäki (2016). Interbank Exposure Networks. *Computational Economics 47*(1), 3–17.

Lopes, H. F. and C. M. Carvalho (2007). Factor Stochastic Volatility with Time Varying Loadings and Markov

Switching Regimes. *Journal of Statistical Planning and Inference 137*(10), 3082–3091.
Minoiu, C. and J. A. Reyes (2013). A Network Analysis of Global Banking: 1978–2010. *Journal of Financial Stability 9*(2), 168–184.
Moghadam, R. and J. Viñals (2010). Understanding Financial Interconnectedness. Mimeo, International Monetary Fund.
Nakajima, J. and M. West (2013). Bayesian Analysis of Latent Threshold Dynamic Models. *Journal of Business and Economic Statistics 31*(2), 151–164.
Roll, R. (1988). R Squared. *Journal of Finance 43*(3), 541–566.
Rothman, A. J., E. Levina, and J. Zhu (2009). Generalized Thresholding of Large Covariance Matrices. *Journal of the American Statistical Association 104*(485), 177–186.
Sarkar, P. and A. W. Moore (2005). Dynamic Social Network Analysis Using Latent Space Models. *ACM SIGKDD Explorations Newsletter 7*(2), 31–40.
Sewell, D. K. and Y. Chen (2015). Latent Space Models for Dynamic Networks. *Journal of the American Statistical Association 110*(512), 1646–1657.
Silva, R. and Z. Ghahramani (2009). The Hidden Life of Latent Variables: Bayesian Learning with Mixed Graph Models. *Journal of Machine Learning Research 10*(Jun), 1187–1238.
Tang, C., M. M. Dungey, M. V. Martin, M. B. González-Hermosillo, and M. R. Fry (2010). Are Financial Crises Alike? Working Paper 10–14, International Monetary Fund.
Tibshirani, R. (1996). Regression Shrinkage and Selection via the LASSO. *Journal of the Royal Statistical Society. Series B 58*(1), 267–288.
Wang, H. (2015). Scaling It Up: Stochastic Search Structure Learning in Graphical Models. *Bayesian Analysis 10*(2), 351–377.
Wang, H. and S. Z. Li (2012). Efficient Gaussian Graphical Model Determination Under G-Wishart Prior Distributions. *Electronic Journal of Statistics 6*, 168–198.

Appendix A. Details of Sampling Approach of the Parameters

This section provides a detailed description of the sampling approach of the parameters.

Appendix A.1. Sampling Σ

For $i = 1, \ldots, n$ and $-i = \{1, \ldots, n\} \backslash \{i\}$, we partition Σ, $S_{y|x}$ and V as

$$\Sigma = \begin{pmatrix} \Sigma_{-i} & \sigma_{-i} \\ \sigma'_{-i} & \sigma_{ii} \end{pmatrix}, \quad S_{y|x} = \begin{pmatrix} S_{-i} & s_{-i} \\ s'_{-i} & s_{ii} \end{pmatrix}, \quad V = \begin{pmatrix} V_{-i} & v_{-i} \\ v'_{-i} & v_{ii} \end{pmatrix} \tag{A.1}$$

where σ_{ii}, s_{ii} and v_{ii} are the i-th diagonal elements of Σ, $S_{y|x}$ and V respectively, σ_{-i}, s_{-i} and v_{-i} are $(n-1) \times 1$ vectors, i.e., the rest of the elements on the i-th column of Σ, $S_{y|x}$ and V, and Σ_{-i}, S_{-i} and V_{-i} are $(n-1) \times (n-1)$ matrices.

From the marginal likelihood function $P(Y|\Sigma)$ in (2.10) and the priors $P(\Sigma|G)$ in (2.14) and $P(G|Z)$ in (2.15), we obtain the following expression

$$P(\Sigma|Y, G) \;\propto\; |\Sigma|^{-\frac{T}{2}} \text{etr}\Big(-\frac{1}{2}\Sigma^{-1} S_{y|x} \Big) \prod_{i \neq j} \exp\Big(-\frac{1}{2} V_{ij}^{-1} \Sigma_{i,j}^2 \Big) \prod_{i=1}^{n} \exp\Big(-\frac{1}{2} V_{ii} \Sigma_{i,i} \Big) \;. \tag{A.2}$$

Using the Sherman-Morrison-Woodbury formula, the inverse of the partitioned Σ is given by

$$\Sigma^{-1} = \begin{pmatrix} \Sigma_{-i}^{-1} + \Sigma_{-i}^{-1} \sigma_{-i} \gamma^{-1} \sigma'_{-i} \Sigma_{-i}^{-1} & -\Sigma_{-i}^{-1} \sigma_{-i} \gamma^{-1} \\ -\sigma'_{-i} \Sigma_{-i}^{-1} \gamma^{-1} & \gamma^{-1} \end{pmatrix} \tag{A.3}$$

where $\gamma = \sigma_{ii} - \sigma'_{-i} \Sigma_{-i}^{-1} \sigma_{-i}$. The determinant of the partitioned Σ is

$$|\Sigma| = |\sigma_{ii} - \sigma'_{-i} \Sigma_{-i}^{-1} \sigma_{-i}| |\Sigma_{-i}| = \gamma |\Sigma_{-i}| \;. \tag{A.4}$$

Following Wang (2015), we consider block updates of Σ by focusing on a column and row at a time. From (A.2) and the partitions in (A.1), the distribution of the elements of the i-th column in Σ, i.e. $(\sigma_{-i}, \sigma_{ii})$, conditional on Y, Σ_{-i}, G is given by

$$P(\sigma_{-i}, \sigma_{ii}|Y, \Sigma_{-i}, G) \propto \gamma^{\frac{T}{2}} \exp\Big(-\frac{1}{2}\Big[\sigma'_{-i}\Sigma_{-i}^{-1}S_{-i}\Sigma_{-i}^{-1}\sigma_{-i}\gamma^{-1} - 2s'_{-i}\Sigma_{-i}^{-1}\sigma_{-i}\gamma^{-1} + s_{ii}\gamma^{-1} + \sigma'_{-i}(v_{ii}\Sigma_{-i}^{-1} + D_v^{-1})\sigma_{-i} + v_{ii}\gamma\Big]\Big) \tag{A.5}$$

where $D_v = \text{diag}(v_{-i})$. We consider a change of variable $(\sigma_{-i}, \sigma_{ii}) \to (\mu, \gamma)$, where $\mu = \sigma_{-i}$ and $\gamma = \sigma_{ii} - \sigma'_{-i}\Sigma_{-i}^{-1}\sigma_{-i})$. The associated Jacobian is a constant independent of (μ, γ). Following Proposition 2 of Wang (2015), the conditional distribution of μ and γ given Y, Σ_{-i}, G is a Gaussian-generalized inverse Gaussian distribution such that

$$\begin{aligned}\mu \mid Y, \Sigma_{-i}, G, \gamma &\sim \mathcal{N}\Big(W^{-1}\Sigma_{-i}^{-1}s_{-i}\gamma^{-1},\ W^{-1}\Big)\\ \gamma \mid Y, \Sigma_{-i}, G, \mu &\sim GIG(q, a, b)\end{aligned}$$

where $W = \Sigma_{-i}^{-1}S_{-i}\Sigma_{-i}^{-1}\gamma^{-1} + v_{ii}\Sigma_{-i}^{-1} + D_v^{-1}$, and (q, a, b) are the parameters of the generalized inverse Gaussian distribution (GIG), where $q = 1 - \frac{1}{2}T$, $a = v_{ii}$, and $b = \mu'\Sigma_{-i}^{-1}S_{-i}\Sigma_{-i}^{-1}\mu - 2s'_{-i}\Sigma_{-i}^{-1}\mu + s_{ii}$. The density of the GIG is given by

$$P(x|q, a, b) = \Big(\frac{a}{b}\Big)^{q/2}\frac{x^{q-1}}{2K_q(\sqrt{ab})}\exp\Big(-\frac{1}{2}[ax + b/x]\Big) \tag{A.6}$$

where K_q is the modified Bessel function of the second kind.

Appendix A.2. Sampling G

Combining $P(\Sigma|G)$ in (2.14) and $P(G|U, \Lambda, \theta)$ in (2.15), the conditional distribution of each edge G_{ij} given Σ, U, Λ, and θ is independent Bernoulli distributed,

$$G_{ij}|\Sigma, U, \Lambda, \theta \sim \text{Ber}\left(\frac{b_{ij1}}{b_{ij1} + b_{ij2}}\right)$$

where $b_{ij1} = \Gamma_{ij}/v_1 \exp\{-\sigma_{ij}^2/(2v_1^2)\}$, and $b_{ij2} = (1 - \Gamma_{ij})/v_0 \exp\{-\sigma_{ij}^2/(2v_0^2)\}$.

Appendix A.3. Sampling Z

Since $Z_{ij}|U, \Lambda, \theta \sim \mathcal{N}(\theta + (U\Lambda U')_{ij}, 1)$ independently and $G_{ij} = \mathbf{1}(Z_{ij} > 0)$, we have:

$$Z_{ij}|G, U, \Lambda, \theta \sim \mathcal{N}(\theta + (U\Lambda U')_{ij}, 1)\ \mathbf{1}(Z_{ij} > 0)^{G_{ij}}\ \mathbf{1}(Z_{ij} < 0)^{1-G_{ij}} \tag{A.7}$$

that is, each Z_{ij} is independently distributed as a truncated version of the prior but conditional on being positive or negatively truncated given G_{ij}.

Appendix A.4. Sampling θ

Following the distribution of Z in (A.7) and a normal prior distribution on θ, the conditional distribution of θ given $\{Z, U, \Lambda\}$ is as follows:

$$P(\theta|Z,U,\Lambda) \;\propto\; \text{etr}\Big(-\frac{1}{4}Z_\theta' Z_\theta\Big)\,\text{etr}\Big(\frac{1}{2}Z_\theta' U\Lambda U'\Big)\;\exp\Big(-\frac{1}{2\tau_\theta^2}(\theta-\theta_0)^2\Big)\;. \tag{A.8}$$

where $Z_\theta = Z - \theta\mathbf{1}\mathbf{1}'$. The posterior distribution of

$$\theta|Z,U,\Lambda \;\sim\; \mathcal{N}\Big(\frac{2\tau_\theta^2}{2+n(n-1)\tau_\theta^2}\Big(\sum_{j>i}(Z-U\Lambda U')_{ij}+\frac{\theta_0}{\tau_\theta^2}\Big),\;\frac{2\tau_\theta^2}{2+n(n-1)\tau_\theta^2}\Big)$$

where $\sum_{j>i}(Z-U\Lambda U')_{ij}$ is a summation of the upper off-diagonals of $(Z-U\Lambda U')$.

Appendix A.5. Sampling Λ

Combining the distribution of Z in (A.7) and the prior distribution of $\Lambda = \text{diag}(\lambda_1, \lambda_2)$ in (2.20), the conditional distribution of Λ given $\{Z, \theta, U\}$ is as follows:

$$P(\Lambda|Z,\theta,U) \;\propto\; \text{etr}\Big(-\frac{1}{2}\Big[\frac{1}{2}\Lambda^2 - Z_\theta' U\Lambda U'\Big]\Big)\prod_{r=1}^{2}\exp\Big(-\frac{1}{2}\frac{\lambda_r^2}{\tau_\lambda^2}\Big)\;. \tag{A.9}$$

Let U_r be the r-th column of U. The posterior distribution for $\lambda_r,\ r = 1, 2$ is given by

$$\lambda_r|Z,\theta,U \;\sim\; \mathcal{N}\Big(\frac{\tau_\lambda^2}{2+\tau_\lambda^2}U_r' Z_\theta U_r,\;\frac{2\tau_\lambda^2}{2+\tau_\lambda^2}\Big)\;.$$

Appendix A.6. Sampling U

Following standard practice, the distribution of U given $\{Z, \theta, \Lambda\}$ is:

$$P(U|\theta,Z,\Lambda) \approx P(U|Z_\theta,\Lambda) \;\propto\; \text{etr}\Big(\frac{1}{2}Z_\theta' U\Lambda U'\Big) \;=\; \text{etr}\Big(\frac{1}{2}\Lambda U' Z_\theta U\Big)\;. \tag{A.10}$$

Comparing (A.10) and (18), we set $C = Z_\theta/2$, and $D = \Lambda$. We sample the columns of U following Gibbs sampler in Hoff (2009). Let $U_{-j} = U\backslash U_j$ denote U excluding the j-th column. For a random draw of $r \sim \{\underline{r}, \bar{r}\}$, we perform the following:

1. obtain N_{-r}, the null space of U_{-r} and compute $x = N_{-r}' U_r$.
2. compute $\tilde{C} = B_u(r,r) N_{-r}' C N_{-r}$
3. update $x \sim P(x|\tilde{C}) \propto \exp(x'\tilde{C}x)$
4. set $U_r = N_{-r}x$

Appendix B. Convergence and Mixing of MCMC

This section provides the results of the mixing of the Markov chains generated by the Gibbs sampler for the samples of θ, Λ, G, and Z in our simulation experiment.

Appendix C. Financial Interconnectedness: Sensitivity Analysis

Table C.4 gives the data description used for our financial application.

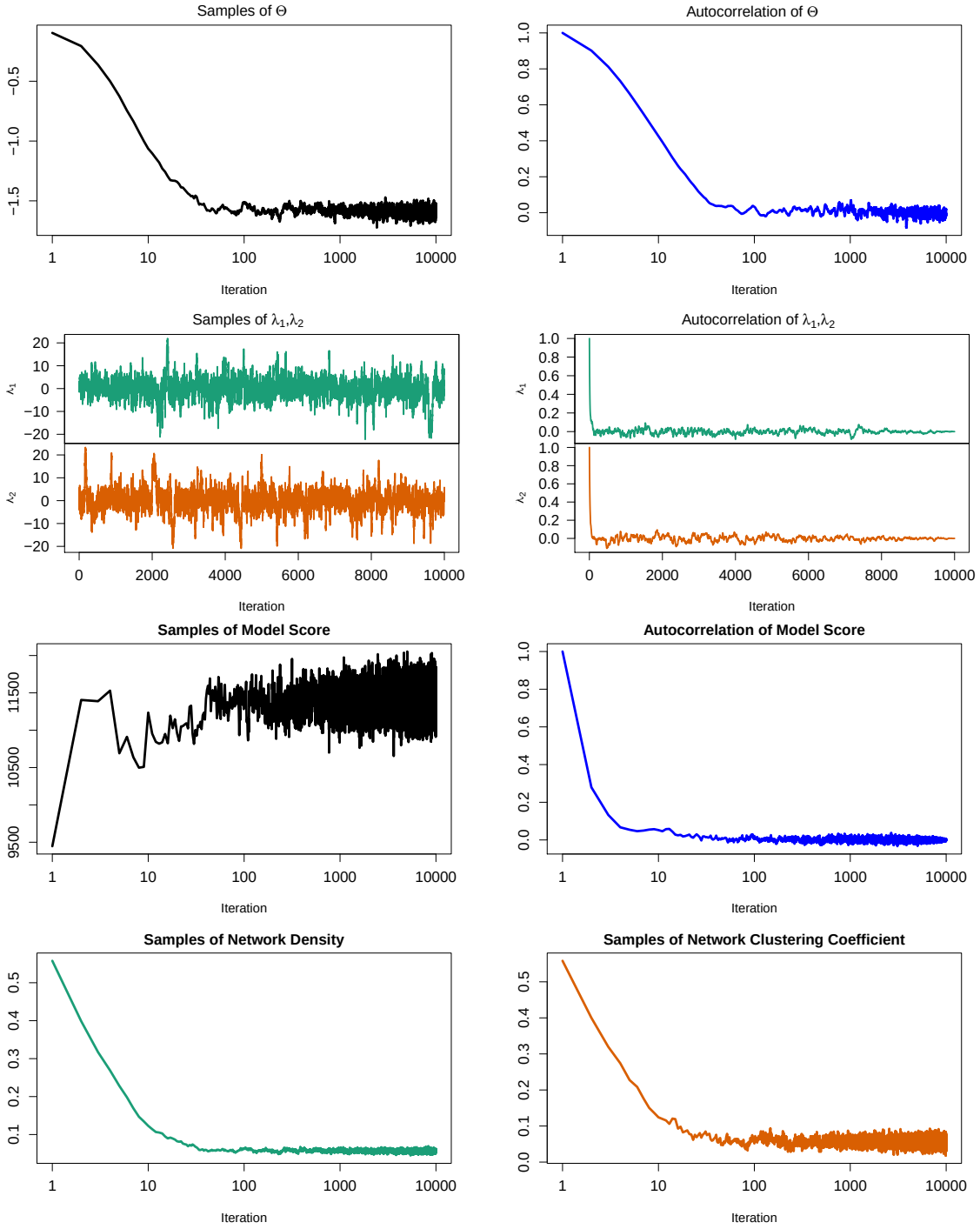

Figure B.10: Plot of samples and autocorrelation of θ, Λ, model score, autocorrelation of model score, the network density and clustering coefficient.

Appendix C.1. Alternative Volatilty Measure

We perform the first robustness check by using volatility of daily returns proxied by a GARCH(1,1) process. The results (see Figure C.11), however, confirm our findings that the interconnectedness and vulnerability of the system that began to manifest early-2008 actually started mounting up from early-2007. The difference between Figure C.11 and Figure 3 is that the SSSL in the former is not smooth and flat as depicted in the latter.

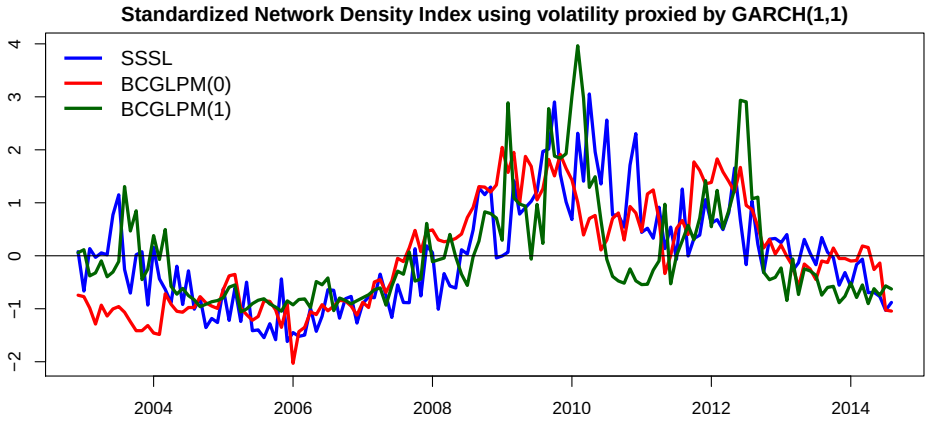

Figure C.11: Standardized network density of the competing methods using volatility proxied by GARCH(1,1).

Appendix C.2. Sensitivity to Spike-and-Slab Parameter

Figure C.12 presents the standardized network density of the methods when the spike-and-slab parameter $v_0 = 0.05$. Compared to Figure 3, the result shows that the SSSL is highly sensitive to the choice of hyperparameters. The BCGLPM methods, on the other hand, are more robust to different specification of the spike-and-slab parameters due to the fact that the method allows for update of the graph priors using information from the latent positions underlying the network.

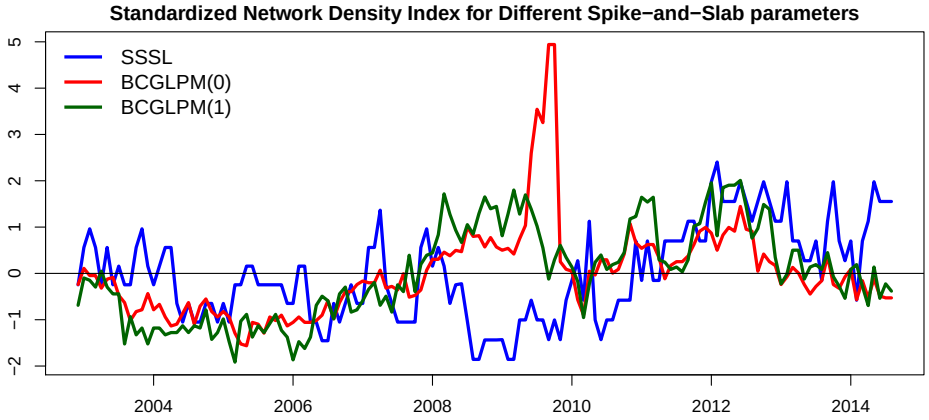

Figure C.12: Standardized network density when the spike-and-slab parameter $v_0 = 0.05$.

Appendix C.3. Robustness to Alternative Rolling Window Size

Figure C.13 shows the standardized network density from a 2-year rolling window estimation. The result is not much different from that of Figure 3 in the sense that the SSSL shows a smooth and flat density between 2004–2007 while the BCGLPM (1) again shows a steady rise over the sub-period.

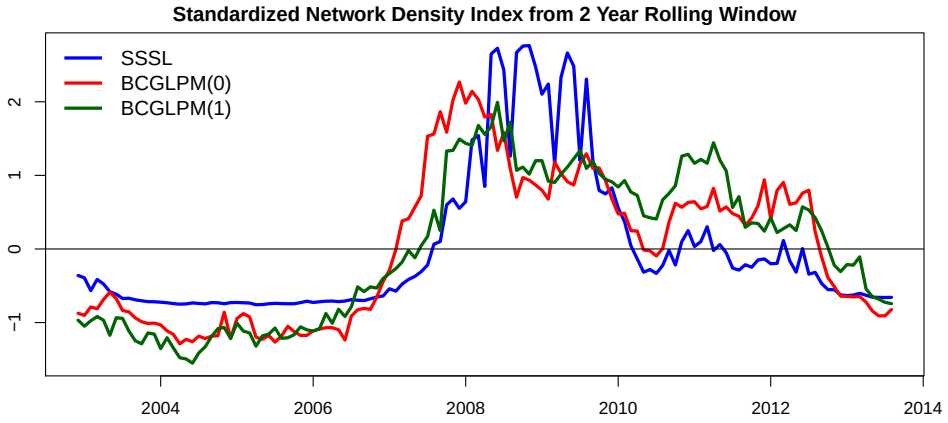

Figure C.13: Standardized network density of the competing methods using a 2-year rolling window.

Appendix C.4. Sensitivity to Alternative Latent Dimensional Space

The results of Figure C.14 shows the standardized network density and similarity of the latent positions from a 3-dimensional latent space model is not different from that of the 2-dimensional latent space reported in Figure 3 and Figure 7.

Table C.4: Financial Data Description By Country and Industry.

No.	Institution	Ticker	Country/Region	Industry
1	S&P 500	GSPC	US	Market Index
2	Dow Jones	DJI	US	Market Index
3	Euro Stoxx 600	DJSTOXX	EU	Market Index
4	Hang Seng Index	HSI	Asia	Market Index
5	Nasdag Composite	IXIC	US	Market Index
6	Euro Stoxx 50	STOXX50E	EU	Market Index
7	Immofiz	IIA	Austria	Real Estate
8	Vienna Insurance Group A	VIG	Austria	Insurance
9	Cofinimmo	COFB	Belgium	Real Estate
10	Credit Suisse Group N	CSGN	Switzerland	Bank
11	Helvetia Holding N	HELN	Switzerland	Insurance
12	PSP Swiss Property Ag	PSPN	Switzerland	Real Estate
13	Swiss Life Holding	SLHN	Switzerland	Insurance
14	Swiss Prime Site	SPSN	Switzerland	Real Estate
15	Swiss Re	SREN	Switzerland	Insurance
16	UBS	UBSN	Switzerland	Bank
17	Zurich Insurance Group	ZURN	Switzerland	Insurance
18	Allianz	ALV	Germany	Insurance
19	Commerzbank	CBK	Germany	Bank
20	Deutsche Bank	DBK	Germany	Bank

No.	Institution	Ticker	Region	Industry
21	Hannover Ruck.	HNR	Germany	Insurance
22	Muenchener Ruck.	MUV	Germany	Insurance
23	Danske Bank	DANSKE	Denmark	Bank
24	Topdanmark	TOP	Denmark	Insurance
25	BBV Argentaria	BBVA	Spain	Bank
26	Mapfre	MAP	Spain	Insurance
27	Banco Santander	SAN	Spain	Bank
28	Sampo A	SAMPO	Finland	Insurance
29	Credit Agricole	ACA	France	Bank
30	BNP Paribas	BNP	France	Bank
31	CNP Assurances	CNP	France	Insurance
32	AXA	CS	France	Insurance
33	Fonciere Des Regions	FDR	France	Real Estate
34	Gecina	GFC	France	Real Estate
35	Societe Generale	GLE	France	Bank
36	ICADE	ICAD	France	Real Estate
37	Natixis	KN	France	Bank
38	Klepierre	LI	France	Real Estate
39	Scor SE	SCR	France	Insurance
40	National Bank of Greece	ETE	Greece	Bank
41	Piraeus Bank	TPEIR	Greece	Bank
42	Bank of Ireland	BIR	Ireland	Bank
43	Banca Monte Dei Paschi	BMPS	Italy	Bank
44	Assicurazioni Generali	G	Italy	Insurance
45	Intesa Sanpaolo	ISP	Italy	Bank
46	Unicredit	UCG	Italy	Bank
47	Unipolsai	US	Italy	Insurance
48	Aegon	AGN	Netherlands	Insurance
49	ING Groep	INGA	Netherlands	Bank
50	Unibail-Rodamco	UL	Netherlands	Real Estate
51	Wereldhave	WHA	Netherlands	Real Estate
52	DNB	DNB	Norway	Bank
53	Storebrand	STB	Norway	Insurance
54	Castellum	CAST	Sweden	Real Estate
55	JM	JM	Sweden	Real Estate
56	Nordea Bank	NDA	Sweden	Bank
57	Aviva	AV	UK	Insurance
58	Barclays	BARC	UK	Bank
59	British Land	BLND	UK	Real Estate
60	Derwent London	DLN	UK	Real Estate
61	Great Portland Estates	GPOR	UK	Real Estate
62	Hammerson	HMSO	UK	Real Estate
63	HSBC Hdg.	HSBA	UK	Bank
64	Hiscox	HSX	UK	Insurance
65	Intu Properties	INTU	UK	Real Estate
66	Land Securities Group	LAND	UK	Real Estate
67	Legal & General	LGEN	UK	Insurance
68	Lloyds Banking Group	LLOY	UK	Bank
69	Old Mutual	OML	UK	Insurance
70	Prudential	PRU	UK	Insurance
71	Royal Bank Of Sctl.Gp.	RBS	UK	Bank
72	RSA Insurance Group	RSA	UK	Insurance
73	Segro	SGRO	UK	Real Estate
74	Shaftesbury	SHB	UK	Real Estate
75	Standard Chartered	STAN	UK	Bank
76	St.Jamess Place	STJ	UK	Insurance

No.	Institution	Ticker	Region	Industry
77	Aflac	AFL	US	Insurance
78	American Intl.Gp.	AIG	US	Insurance
79	Arthur J Gallagher	AJG	US	Insurance
80	Allstate	ALL	US	Insurance
81	AON Class A	AON	US	Insurance
82	American Express	AXP	US	Bank
83	Bank of America	BAC	US	Bank
84	BB&T	BBT	US	Bank
85	Bank of New York Mellon	BK	US	Bank
86	BOK Finl.	BOKF	US	Bank
87	Berkshire Hathaway A	BRKA	US	Insurance
88	Brown & Brown	BRO	US	Insurance
89	Citigroup	C	US	Bank
90	Chubb	CB	US	Insurance
91	Comerica	CMA	US	Bank
92	CNA Financial	CNA	US	Insurance
93	Capital One Finl.	COF	US	Bank
94	Corrections Amer New	CXW	US	Real Estate
95	Duke Realty Corporation	DRE	US	Real Estate
96	Equity Lifestyle Props.	ELS	US	Real Estate
97	Essex Property Tst.	ESS	US	Real Estate
98	Fifth Third Bancorp	FITB	US	Bank
99	Federal Realty Inv.Tst.	FRT	US	Real Estate
100	General Gw.Props.	GGP	US	Real Estate
101	Goldman Sachs Gp.	GS	US	Bank
102	Huntington Bcsh.	HBAN	US	Bank
103	Hudson City Banc.	HCBK	US	Bank
104	HCC Insurance Hdg.	HCC	US	Insurance
105	Welltower - Health Care Reit	HCN	US	Real Estate
106	HCP	HCP	US	Real Estate
107	Hartford Finl.Svs.Gp.	HIG	US	Insurance
108	Host Hotels & Resorts	HST	US	Real Estate
109	JP Morgan	JPM	US	Bank
110	Keycorp	KEY	US	Bank
111	Kimco Realty	KIM	US	Real Estate
112	Loews	L	US	Insurance
113	Lincoln National	LNC	US	Insurance
114	Liberty Property Tst.	LPT	US	Real Estate
115	Macerich	MAC	US	Real Estate
116	Metlife	MET	US	Insurance
117	Markel	MKL	US	Insurance
118	Marsh & Mclen	MMC	US	Insurance
119	Morgan Stanley	MS	US	Bank
120	M&T Bank	MTB	US	Bank
121	Mitsub.Ufj Finl.Gp. Adr	MTU	US	Insurance
122	Northern Trust	NTRS	US	Bank
123	Realty Income	O	US	Real Estate
124	Principal Finl.Gp.	PFG	US	Insurance
125	Progressive Ohio	PGR	US	Insurance
126	Prologis	PLD	US	Real Estate
127	PNC Finl.Svs.Gp.	PNC	US	Bank
128	Prudential Finl.	PRU	US	Insurance
129	Public Storage	PSA	US	Real Estate
130	Regency Centers	REG	US	Real Estate
131	Regions Finl.New	RF	US	Bank
132	Charles Schwab	SCHW	US	Bank

No.	Institution	Ticker	Region	Industry
133	Sl Green Realty	SLG	US	Real Estate
134	Simon Property Group	SPG	US	Real Estate
135	Suntrust Banks	STI	US	Bank
136	State Street	STT	US	Bank
137	Torchmark	TMK	US	Insurance
138	Travelers Cos.	TRV	US	Insurance
139	Unum Group	UNM	US	Insurance
140	US Bancorp	USB	US	Bank
141	Vornado Realty Trust	VNO	US	Real Estate
142	Ventas	VTR	US	Real Estate
143	Wells Fargo & Co	WFC	US	Bank
144	W R Berkley	WRB	US	Insurance
145	Alleghany	Y	US	Insurance
146	Zions Bancorp.	ZION	US	Bank
147	Bear Stearns	BSC	US	Bank
148	Countrywide Financial Corp.	CCR	US	Bank
149	Lehman Brothers	LEHM	US	Bank
150	Merrill Lynch	MER	US	Bank

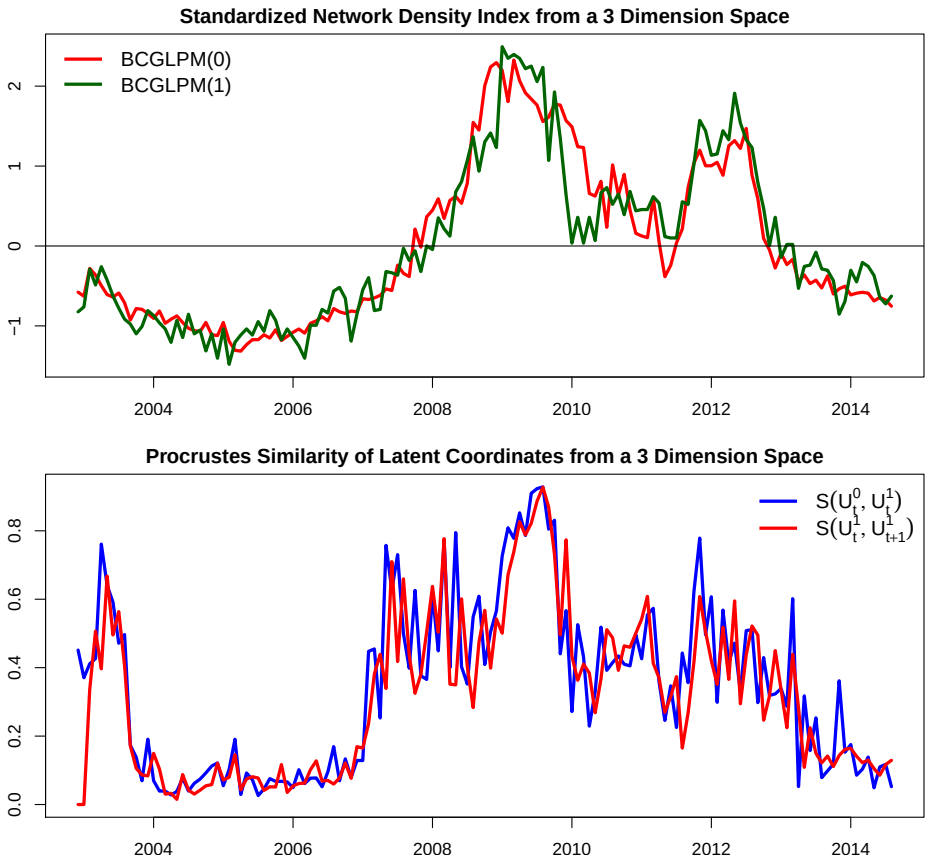

Figure C.14: Standardized network density and similarity of latent positions from a 3-dimensional latent space.