# Modeling Systemic Risk with Markov Switching Graphical SUR Models*

Daniele Bianchi† Monica Billio‡ Roberto Casarin§ Massimo Guidolin¶

**Abstract**

We propose a Markov Switching Graphical Seemingly Unrelated Regression (MS-GSUR) model to investigate time-varying systemic risk based on a range of multi-factor asset pricing models. Methodologically, we develop a Markov Chain Monte Carlo (MCMC) scheme in which latent states are identified on the basis of a novel weighted eigenvector centrality measure. An empirical application to the S&P100 constituents shows that cross-firm connectivity significantly increased over the period 1999-2003 and the financial crisis of 2008-2009. Finally, we provide evidence that firm-level centrality does not correlate with market values and is instead positively linked to realized financial losses.

**Keywords:** Markov Regime-Switching, Weighted Eigenvector Centrality, Graphical Models, MCMC, Systemic Risk, Network Connectivity

**JEL codes:** C11, C15, C32, C58

*We are grateful to Luc Bawuens, Guido Consonni, Fulvio Corsi, Francis X. Diebold, Sylvia Frühwirth-Schnatter, Siam J. Koopman, Fabrizio Lillo, Chris Sims, Mike West, Kamil Yilmaz, for their helpful comments and suggestions. We also thank seminar participants at the Warwick Business School, Scuola Normale Superiore of Pisa, HSE of Moscow, the NBER Summer Institute 2015, the NBER-NSF Time Series conference 2015, the 8th Annual SoFiE conference, the SYRTO Conference on Systemic Risk, the 2nd Vienna Workshop on High-Dimensional Time Series in Macroeconomics and Finance, the European Seminar on Bayesian Econometrics ESOBE 2015, and the FMA Orlando 2015. Monica Billio and Roberto Casarin acknowledge financial support from the European Union, Seventh Framework Programme FP7/2007-2013 under grant agreement SYRTO-SSH-2012-320270, from the Institut Europlace de Finance under the *Systemic Risk* grant, and from the Italian Ministry of Education, University and Research (MIUR) PRIN 2010-11 grant MISURA. This research used the SCSCF multiprocessor cluster system at University Ca' Foscari of Venice. A previous version of the paper was circulating under the title "Modeling Contagion and Systemic Risk".

†Warwick Business School, University of Warwick, Coventry, UK. Daniele.Bianchi@wbs.ac.uk

‡Department of Economics, University Ca' Foscari of Venice, Venice, Italy. billio@unive.it

§Department of Economics, University Ca' Foscari of Venice, Venice, Italy. r.casarin@unive.it

¶Department of Finance, Baffi-CAREFIN and IGIER, Bocconi University, Milan, Italy. massimo.guidolin@unibocconi.it

# 1 Introduction

The 2008-2009 financial crisis has shown that liquidity and valuation shocks may quickly propagate across the economic system and affect financial institutions operating in different markets, with different size and business structure, thus causing widespread losses and domino effects. Therefore, understanding the dynamics of cross-asset and cross-equity linkages is of key importance to both systemic risk management purposes and to forecast and deal with contagion waves in times of crisis. Systemic risk shocks are conventionally referred in the empirical network literature as abrupt increases in the density of cross-firm connectivity (see, e.g. Billio, Getmansky, Lo, and Pelizzon 2012, and references therein). Modeling firms connectedness has a long history in the empirical finance literature, starting from the seminal contributions by Cont and Bouchaud (2000) and Markose (2005). Recently, network analysis has been further supported by a series of papers that have shown its in- and out-of-sample superior performance over traditional, correlation-based approaches.[1]

We extend the existing literature in a number of ways. First, we propose a system-wide inferential scheme based on a Markov-switching graphical model, that allows to simultaneously consider all of the possible linkages among firms through constraints on the regime-specific conditional dependence structure. Second, we propose an identification scheme for different regimes of cross-firm connectivity based on a novel weighted eigenvector centrality measure, which is related to both the number and the weight of connections between graph vertices. Third, we provide an asset pricing application based on otherwise standard multi-factor pricing models in which the exposures of the assets to risk factors (betas) are allowed to change according to the regimes in cross-firm connectivity. This allows us to develop a unified framework where *systematic* and *systemic* risks are not mutually exclusive, in the sense that firm-specific exposures to sources of systematic risk may directly depend on the level of

[1] See, e.g. Forbes and Rigobon (2000), Forbes and Rigobon (2002), Corsetti, Pericoli, and Sbracia (2005), Corsetti, Pericoli, and Sbracia (2011), Billio et al. (2012), Hautsch, Schaumburg, and Schienle (2015), Barigozzi and Brownlees (2014), Diebold and Yilmaz (2014),Timmermann, Blake, Tonks, and Rossi (2014), Brownlees, Nualart, and Sun (2014), and Diebold and Yilmaz (2015), among others.

aggregate network connectivity.[2] Finally, we provide a Metropolis-within-Gibbs sampling scheme which allows to jointly draw both the parameters of the factor pricing model, the latent states and the underlying regime-specific graphs.

Methodologically, we build upon the Gaussian Graphical model for multi-variate systems proposed by Whittaker (1990), Dawid and Lauritzen (1993), Lauritzen (1996), Carvalho and West (2007), Wang and West (2009), Wang (2010), Rodriguez, Lenkoski, and Dobra (2011), Wang, Reeson, and Carvalho (2011) and Ahelegbey, Billio, and Casarin (2016). In particular Wang et al. (2011) developed a dynamic matrix-variate graphical model which allows to capture conditional dependencies under a time-invariant graphs. We generalize and extend their framework by introducing Markov-switching dynamics in the graph structure within a Seemingly Unrelated Regression (SUR) model. In particular, we propose a new Markov Switching Graphical SUR (MS-GSUR) which allows to identify different regimes of network connectivity (systemic risk). The regime-switching identification problem is solved based on the graph-theoretic properties of the state-specific conditional dependence structures of the model error terms. More specifically, we propose a new weighted eigenvector centrality measure, which accounts not only for the number of adjacent nodes, but also for the weights of the edges and for the number of indirect connections between nodes, i.e. walks between nodes. We show formally that our measure can be interpreted as a weighted sum over the walks and can be related to other measures, such as Bonacich (1972)'s and communicability (see e.g., Estrada and Hatano 2008 and Estrada and Hatano 2009), used in the analysis of the topological features of complex networks.

The empirical application of our MS-GSUR model focuses on the cross-section of daily excess returns of the S&P100 Index constituents over the period 1996-2014. The emphasis on stock returns is motivated by a widespread desire of policy makers and regulators to incorporate the most current information for the purpose of systemic risk measurement: stock prices of largely traded stocks reflect information more rapidly

[2]In this paper, we define systemic risk as the risk caused by the possibility that a firm-specific event may be severe enough to cause pervasive instability in the whole economic system; we define instead systematic risk as the risk inherent in aggregate market and macroeconomic conditions that cannot be simply diversified away. Moreover, we use the notions of network connectivity, connectedness, and of systemic risk interchangeably.

than other non-traded measures such as accounting variables, the volumes of deposits, or loans, even though the mechanisms involving the latter variables have been at the epicenter of a number of financial crises. The cross-firm connectivity is estimated on the basis of the residual covariance structure of stock returns, conditioning on a set tradable risk factors used in some of the most popular factor models, namely the CAPM, the three-factor model of Fama and French (1993), and the Merton (1973) intertemporal CAPM (I-CAPM). Interestingly, our results are robust across model specifications.

The main empirical findings show that the dynamics of systemic risk can be captured by two regimes, in which a state of high connectedness characterizes the period 1999-2003 (marked by the passing of the Gramm-Leach-Bliley act, the inflating and burst of the dot.com bubble, and the ensuing financial scandals) and subsequently, the great financial crisis of 2008-2009. We show that a few financial institutions turned out to heavily outweigh other firms in the network during such time intervals and that shocks to the financial sector turned out to be the most systemically important. Finally, both a simple cross-sectional regression and rank-correlation analysis show that market capitalization does not sensibly drive the relevance of a given firm within the network. However, firms which are more relevant within the network are more likely to suffer significant losses during periods of high systemic risk.

The remainder of the paper is organized as follows. Section 2 and 3 describe the modeling framework and the estimation strategy. Section 4 presents the empirical results from our application. Section 5 concludes.

# 2 A Markov Switching Graphical SUR Model

Let $y_{it}$ be the stock returns of the $ith$ firm in excess of the risk-free rate at time $t$, and $\mathbf{x}_{it}$ the $m_i$-dimensional vector of systematic risk factors. In its basic formulation, each stock return time series is modeled as a dynamic multi-factor linear model

$$y_{it} = \mathbf{z}_{it}'\boldsymbol{\beta}_i(s_t) + \varepsilon_{it}, \qquad t = 1, \ldots, T, \qquad i = 1, \ldots, n \tag{1}$$

where the matrix $\mathbf{z}_{it} = (1, \mathbf{x}'_{it})'$ includes an intercept plus the $m_i$ covariates, $\boldsymbol{\beta}_i(s_t) = (\beta_{i0}(s_t), \beta_{i1}(s_t), \ldots, \beta_{im_i}(s_t))'$ is a $(m_i + 1)$-vector of time-varying regression coefficients, and $\varepsilon_{it}$ is an error term that can be identified with a firm-specific idiosyncratic risk factor when $Cov(\mathbf{z}_{it}, \varepsilon_{it}) = \mathbf{0}$. The multi-factor pricing model in (1) is fairly general because it represents a reduced-form approximation of a linear pricing kernel (see, e.g. Liew and Vassalou 2000, Cochrane 2001, and Vassalou 2003). The model can be re-written in a more compact form as a SUR, i.e.

$$\mathbf{y}_t = Z'_t \boldsymbol{\beta}(s_t) + \boldsymbol{\varepsilon}_t, \qquad t = 1, \ldots, T \tag{2}$$

with $\mathbf{y}_t = (y_{1t}, \ldots, y_{nt})'$ the dependent variable vector, $Z_t = \text{diag}\{\mathbf{z}_{1t}, \ldots, \mathbf{z}_{nt}\}$ the block-diagonal matrix of covariates, $\boldsymbol{\beta}(s_t)' = (\boldsymbol{\beta}'_1(s_t), \ldots, \boldsymbol{\beta}'_n(s_t))$ the time-varying coefficient vector of dimension $m = n + \sum_{i=1}^{n} m_i$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ the error terms vector. We assume that risk factors are common across stocks and the error terms have a full time-varying variance-covariance matrix, but are independently and normally distributed conditionally on the latent state $s_t$, $t = 1, \ldots, T$, i.e. $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \Sigma(s_t))$.

In our model, the time-varying network is identified by the inverse variance-covariance matrix $\Omega(s_t) = \Sigma(s_t)^{-1}$ of the SUR error terms and a set of zero-restrictions implied by an underlying graph $G(s_t) \in \mathcal{G}$ where $\mathcal{G}$ is the space of undirected graphs. More precisely, we introduce a state-dependent graph defined by the ordered pair of disjoint sets $G(s_t) = (V(s_t), D(s_t))$ where $V(s_t)$ is the *vertex set* of $n$ nodes and $D(s_t) \subset V(s_t) \times V(s_t)$ defines the *edge set* in the state $s_t$.[3] The nodes represent the firms within the economy and the edges define the presence of interconnections among and across firms. If two nodes $i \in V(s_t)$ and $j \in V(s_t)$ are adjacent in the graph, i.e. $\{i, j\} \in D(s_t)$, then there is an interconnection between two firms. If $G(s_t)$ is an undirected graph, a Gaussian graphical model for the error terms is defined by the assumption that $\boldsymbol{\varepsilon}_t$ is normally distributed with independent elements implied by

[3]See, e.g. Bollobás (1998, 2001) and for an introduction to graph theory.

$G(s_t)$, i.e.,[4]

$$\varepsilon_{it} \perp \varepsilon_{jt} \,|\varepsilon_{V(s_t)\setminus\{i,j\}} \iff \omega_{ijt} = 0, \tag{3}$$

with $\varepsilon_{V(s_t)\setminus(i,j)} = \{\varepsilon_{lt}; l \in V(s_t), l \neq i, j\}$ $\omega_{ijt}$ $(i,\, j = 1, \ldots, n)$ the $(i,\, j)$-$th$ element of the precision matrix $\Omega(s_t)$. Hence, $\Omega(s_t) \in \mathcal{M}(G(s_t))$ with $\mathcal{M}(G(s_t))$ the set of all positive-definite symmetric matrices with elements equal to zero for all $\{i, j\} \notin D(s_t)$.

The Gaussian nature of our modeling framework might present clear limitations as it does not allow to capture co-dependencies in higher moments. While a distinctive attention to covariances and correlations is typical of a classical finance (i.e., of mean-variance derivation) approach, clearly this overlooks other important aspects. However, given our goal to use standard I-CAPM-style models to separate the notion of systematic risk from systemic risk our choice seems a sensible one as a first step of analysis. Moreover, given the residual nature of systemic risk, we assume there is no given direction in the linkages among firms. This is of course a limitation of our framework and we leave extensions to overcome its undirected nature for future research.

The time-variation in the parameters of the MS-GSUR model is driven by a latent first-order Markov chain process $s_t$ with time-homogeneous transition probabilities $\text{Prob}(s_t = j|s_{t-1} = i) = \pi_{ij}$, $i, j = 1, \ldots, K$, for all $t = 1, \ldots, T$. The choice of a regime-dependent dynamics is motivated by the common definition of contagion and systemic risk as an abrupt increase in the cross-sectional dependence structure of firms/sectors/countries, e.g. Forbes and Rigobon (2000), that a framework with recurrent regimes may easily capture. Also, this class of models is popular in the finance literature as they allow for an intuitive economic interpretation of different market phases (see, e.g., Ang and Bekaert 2002, and Guidolin and Timmermann

[4]See, e.g., Erdös and Rényi (1959), Dempster (1972), Dawid and Lauritzen (1993), Giudici and Green (1999), and Carvalho and West (2007) for more details on static specifications of graphical models.

2008). The Markov-switching dynamics of the SUR coefficients writes as

$$\boldsymbol{\beta}(s_t) = \sum_{k=1}^{K} \boldsymbol{\beta}_k \mathbb{I}_{\{k\}}(s_t) \tag{4}$$

with $\mathbb{I}_{\{k\}}(s_t)$ the indicator function which takes value one when the state $s_t$ takes value $k$ at time $t$, and zero otherwise, and $\beta_k$ the vector of regime-specific betas. We assume that for each state $s_t = k$ there is a regime-specific covariance matrix $\Sigma(s_t)$ constrained by the state-specific graph $G_k = (V_k, D_k)$ such that,

$$\Sigma(s_t) = \sum_{k=1}^{K} \Sigma_k \mathbb{I}_{\{k\}}(s_t) \tag{5}$$

The regime-dependent feature of the covariance structure allows to address potential heteroskedasticity biases which are typical of correlation-based measures, as discussed in Forbes and Rigobon (2002). Also, the topological features of the state-specific graph $G_k$ play a crucial role in the estimation of the model since they allow us to identify regimes of low vs. high systemic risk.

The factor model specification in (2)-(5) implies that systematic and systemic risks are not mutually exclusive. In particular, while the exposures to systematic risks, i.e., the betas, depend on the latent state, the latter, as shown below, is directly identified from the network defined by model residuals, which in their turn depend on the regression betas. A micro-foundation of such relationship is provided in Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012). The key insight is that if cross-firm linkages are sufficiently asymmetric, then firm-specific shocks cannot necessarily be diversified away, but instead may aggregate into macroeconomic fluctuations. Consider for instance a negative supply shock to the oil market, e.g., OPEC's decision to cut production. For a given level of aggregate demand, oil prices will increase, which will cause a decline in the profitability of transportation and manufacturing firms, and will subsequently cause negative effects for the fundamentals of the the consumers' discretionary sector. Thus, a shock originated at the firm or sector level spreads throughout the economy to progressively generate a systematic, aggregate

shock. As a result, firms that are more central in the network will present higher exposure to systematic risks.[5]

# 3 Inference on Parameters and Network Structure

We focus on decomposable graphs $G_k$, $k = 1, \ldots, K$ (see, e.g., Dawid and Lauritzen 1993). If the joint distribution of excess stock returns is Markov with respect to a decomposable graph $G_k$, the joint density of $\mathbf{y}_t$ given $s_t = k$ factorizes as

$$p\left(\mathbf{y}_t | Z_t, \boldsymbol{\beta}_k, \Sigma_k, G_k, s_t\right) = \prod_{p \in \mathcal{P}_k} p\left(\mathbf{y}_{pt} | Z_{pt}, \boldsymbol{\beta}_{pk}, \Sigma_{pk}, s_t\right) / \prod_{b \in \mathcal{B}_k} p\left(\mathbf{y}_{bt} | Z_{bt}, \boldsymbol{\beta}_{bk}, \Sigma_{bk}, s_t\right) \quad (6)$$

where $\mathcal{P}_k$ is the set of complete prime components and $\mathcal{B}_k$ the set of separators. For each subgraph $g \in \{\mathcal{P}_k, \mathcal{B}_k\}$, $\mathbf{y}_{gt} = (y_{it} : i \in g)$, $Z_{gt}$ is the corresponding matrix of covariates, $\boldsymbol{\beta}_{gk} = (\boldsymbol{\beta}_{ik} : i \in g)$ the subset of factor loadings, and $\Sigma_{gk}$ the relative block of the covariance matrix of the residuals from $\Sigma_k$. Each term in (5) has a multivariate normal distribution, $\mathbf{y}_g \sim \mathcal{N}(Z_g' \boldsymbol{\beta}_{gk}, \Sigma_{gk})$, where $\Omega_{gk} = \Sigma_{gk}^{-1}$ is a full positive-definite symmetric matrix. Given the graph $G_k$, the joint distribution (5) is completely defined by the component-specific marginal betas, covariates, and covariance matrices (see, e.g. Giudici and Green 1999 and Carvalho and West 2007).

## 3.1 Prior Specification

Dawid and Lauritzen (1993) define a class of probability distributions for covariance matrices on decomposable graphs called Hyper-inverse Wishart. Conditional on the graph $G_k$, the prior distribution of the $k$-th state covariance matrix $\Sigma_k$ is the Hyper-inverse Wishart

$$\Sigma_k \sim \mathcal{HIW}_{G_k}\left(a_k, A_k\right) \quad (7)$$

[5] In this respect, even though the interpretation of the model is non-structural, modelling the dynamics of propagation of reduced-form shocks may provide a useful tool to understand strength and patterns of systemic risk.

with $a_k$ and $A_k$ the degrees of freedom and the location parameters, respectively. We denote this prior distribution with $p(\Sigma_k|G_k)$. For each state $s_t = k$, the hyper-inverse Wishart represents the unique conjugate "local prior" for complete prime components that are inverse Wishart distributed (see Carvalho, West, and Massam 2007 for a detailed discussion). Let $\mathcal{P}_k = \{P_{1,k}, \ldots, P_{n_P,k}\}$ and $\mathcal{B}_k = \{B_{1,k}, \ldots, B_{n_B,k}\}$ be the set of $n_P, k$ prime components and $n_B, k$ separators, respectively. By generating the tree representation of the prime components, the density of the hyper-inverse Wishart for $\Sigma_k$ conditional on $G_k$ has expression

$$p(\Sigma_k|G_k) = \prod_{j=1}^{n_P,k} p(\Sigma_{P_j,k}) \prod_{i=1}^{n_B,k} p(\Sigma_{B_i,k})^{-1} \tag{8}$$

where for each prime component $\Sigma_{P_j,k} \sim \mathcal{IW}\left(a_k, A_{P_j,k}\right)$ the density is given by

$$p(\Sigma_{P_j,k}) \propto |\Sigma_{P_j,k}|^{-\left(a_k+2T_{P_j,k}\right)/2} \exp\left\{-\frac{1}{2}\text{tr}(\Sigma_{P_j,k}^{-1} A_{P_j,k})\right\} \tag{9}$$

with $T_{P_j,k} = Card(P_j, k)$ the cardinality of $P_j, k$, and $A_{P_j,k}$ the $j$th diagonal block of $A_k$ corresponding to $\Sigma_{P_j,k}$ (see Hammersley and Clifford 1971 and Dempster 1972).[6] The prior over the graph structure is defined as a Bernoulli distribution with parameter $\psi$ on each edge inclusion probability used as an initial sparse-inducing prior. That is, a $n$-node graph $G_k = (V_k, D_k)$ has a prior probability

$$p(G_k) \propto \prod_{i,j} \psi^{d_{ij,k}} (1-\psi)^{(1-d_{ij,k})} \propto \psi^{|D_k|} (1-\psi)^{N-|D_k|} \tag{10}$$

where $d_{ij,k} = 1$ if $\{i, j\} \in D_k$ and 0 otherwise and $|D_k|$ denotes the cardinality of $D_k$, which is the number of edges, or size, of the graph $G_k$ (see Bollobás 1998, ch. 1). Given the graph is undirected, the cardinality of $|D_k|$ is given by $\sum_{i,j \in V_k} d_{ij}/2$, which is equal to $N = n(n-1)/2$ if the $n$-node graph is complete, i.e. there is an edge

[6] An important question for our framework is whether hyper-inverse Wishart priors on graphical models make sense at all in other than very low dimensions. Indeed, priors inherit the fundamental limitation of Inverse-Wishart priors, namely a single dispersion parameter $a_k$ for all and every aspect of uncertainty and dependency among the variance matrix entries. This assumption can be relaxed at the price of further increasing the computational complexity of the model. However, to investigate the trade-off between imposing more complex prior structures on the covariance matrix and the gains in terms of systemic risk modeling is beyond the scope of the paper.

between all pairs of nodes. The prior (10) has its peak at $N\psi$ providing a flexible way to directly control for prior model complexity.[7] To induce sparsity and hence obtain a parsimonious representation of the interdependence structure implied by a graph, we choose $\psi = 2/(n-1)$ which would provide a prior mode in correspondence to $n$ edges. In addition to this baseline specification, we try two alternative prior specifications for $p(G_k)$ which imply either an empty edge set (i.e., an empty graph), or a complete graph , i.e., $d_{ij,k} = 1$ for $i,j = 1, \ldots, n$, $i \neq j$. However, in our Gibbs sampler, the posterior distribution of the set $G_k$ tends to reach a similar posterior median for different starting priors, confirming the robustness in our findings.[8] The prior for the factor betas is chosen to be independent of the covariance structure,

$$\boldsymbol{\beta}_k \sim \mathcal{N}(\mathbf{m}_k, M_k) \tag{11}$$

We choose for the prior distributions for the regression parameters centered around zero and rather uninformative and common across states. The prior distribution for the $k$th row of the transition matrix $\Pi$, i.e., $\boldsymbol{\pi}_k = (\pi_{k1}, \ \ldots, \ \pi_{kK})'$ is a Dirichlet, i.e.

$$\boldsymbol{\pi}_k \sim \mathcal{D}ir(c_{k1}, \ldots, c_{kK}) \tag{12}$$

with $c_{lk}$ the concentration hyper-parameter.

## 3.2 Posterior Approximation

Let us denote with $\mathbf{y}_{\tau:t} = (\mathbf{y}_\tau, \ldots, \mathbf{y}_t)$, $\tau \leq t$, the data between observations $\tau$ and $t$., and with $G = (G_1, \ldots, G_K)$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K)$ the collections of state-specific graphs and parameters, respectively, where $\boldsymbol{\theta}_k = (\boldsymbol{\beta}_k, \Sigma_k, \boldsymbol{\pi}_k)$, $k = 1, \ldots, K$. The

[7] An alternative uniform prior might have been used instead. However as pointed out in Jones, Carvalho, Dobra, Hans, Carter, and West (2005), a uniform prior over the space of all graphs is biased towards a graph with half of the total number of possible edges.

[8] Further results on the properties of our Gibbs sampler are available upon request.

complete likelihood of the data is then defined as

$$p\left(\mathbf{y}_{1:T}, \mathbf{s}_{1:T}|\boldsymbol{\theta}, G\right) = \prod_{t=1}^{T} p\left(\mathbf{y}_t|s_t, \boldsymbol{\theta}, G\right) p\left(s_t|s_{t-1}, \boldsymbol{\theta}\right) p\left(s_0\right) \tag{13}$$

With reference to our application, the marginal likelihood of the data accommodates the existence of fat tails in the distribution of excess asset returns. Indeed, given the local conjugate priors and the hyper-Markov structure of the graph $G_k$, one can show that the marginal distribution conditional on the sequence of states $\mathbf{s}_{1:T}$ is an *hyper Student-t* (see, e.g. Dawid and Lauritzen 1993 for more details).

Let $p(\boldsymbol{\theta}, G) \propto \prod_{k=1}^{K} p(\boldsymbol{\beta}_k)p(\Sigma_k|G_k)p(G_k)$ be the joint prior distribution, then the joint posterior is $p(\boldsymbol{\theta}, G|\mathbf{s}_{1:T}, \mathbf{y}_{1:T}) \propto p\left(\mathbf{y}_{1:T}, \mathbf{s}_{1:T}|\boldsymbol{\theta}, G\right) p(\boldsymbol{\theta}, G)$. Because such distribution is not tractable, the estimator of the parameters and graphs cannot be obtained in analytical form. We approximate the posterior distribution by implementing a multi-move Gibbs sampling algorithm (see, e.g., Roberts and Sahu 1997 and Casella and Robert 2004), where the graph structure, the hidden states, and the parameter are all sampled in blocks.[9] At each iteration the Gibbs sampler sequentially cycles through the following steps:

1. sample $\mathbf{s}_{1:T}$ given the graphs $G$, the parameters $\boldsymbol{\theta}$ and the observations $\mathbf{y}_{1:T}$;
2. sample $\Sigma_k$ given $\mathbf{s}_{1:T}$, $G_k$ and $\mathbf{y}_{1:T}$, for $k = 1, ..., K$
3. sample $G_k$ given $\Sigma_k$, $\mathbf{s}_{1:T}$ and $\mathbf{y}_{1:T}$, for $k = 1, ..., K$
4. sample $\boldsymbol{\beta}_k$ given $\mathbf{s}_{1:T}$, $\Sigma_k$, $G_k$ and $\mathbf{y}_{1:T}$, for $k = 1, ..., K$
5. sample $\Pi$ given $\mathbf{s}_{1:T}$.

We extract the latent states $s_t$, $t = 1, \ldots, T$ by using a forward filtering backward sampling (FFBS) algorithm (see Frühwirth-Schnatter 1994 and Carter and Kohn 1994). Because the latent state is discretely valued, the FFBS is applied in its Hamilton (1994)'s form.

Denote $\mathcal{T}_k = \{t : s_t = k\}$, and $T_k = Card(\mathcal{T}_k)$. Also, let $\mathbf{e}_{tk} = \mathbf{y}_t - Z_t'\boldsymbol{\beta}_k$ be the residuals conditional on the state $s_t = k$ and $A_k^* = \sum_{t\in\mathcal{T}_k} \mathbf{e}_{tk}\mathbf{e}_{tk}'$. Given the local

[9]Such procedure allows to sample from the joint posterior distribution of the latent states and parameters without rely on asymptotic distributional assumptions.

conjugate prior in Eq.(7) the posterior for $\Sigma_k$ is:

$$p\left(\Sigma_k|\mathbf{y}_{1:T},\boldsymbol{\theta},\mathbf{s}_{1:T},\boldsymbol{\beta}_k\right) \propto \mathcal{HIW}_{G_k}\left(a_k+T_k, A_k+A_k^*\right), \tag{14}$$

See Appendix A for a proof. To sample the graphs $G_k, k=1,...,K$ we compute the unnormalized posterior over graphs $p(G_k|\mathbf{y}_{1:T},\mathbf{s}_{1:T}) \propto p(\mathbf{y}_{1:T}|\mathbf{s}_{1:T},G_k)p(G_k)$, for any specified state $k$ (see, e.g. Giudici and Green 1999 and Jones et al. 2005). As in the proof of Eq. (14), given the prior independence assumption of the parameters across regimes,

$$p\left(G_k|\mathbf{y}_{1:T},\mathbf{s}_{1:T}\right) \propto p\left(G_k\right) \int\int p(\mathbf{y}_{\mathcal{T}_k}|\mathbf{s}_{\mathcal{T}_k},\boldsymbol{\beta}_k,\Sigma_k)p(\boldsymbol{\beta}_k)p(\Sigma_k|G_k)d\boldsymbol{\beta}_k d\Sigma_k \tag{15}$$

where

$$p(\mathbf{y}_{\mathcal{T}_k}|\mathbf{s}_{\mathcal{T}_k},\boldsymbol{\beta}_k,\Sigma_k) \propto |\Sigma_k|^{-T_k/2}\exp\left\{-\frac{1}{2}\text{tr}\left(\Sigma_k^{-1}A_k^*\right)\right\}.$$

To evaluate this integral we follow Chib (1995) and Wang (2010) and approximate the marginal likelihood via a local-move Metropolis-Hastings step based on the conditional posterior $p(G_k|\mathbf{y}_{1:T},\mathbf{s}_{1:T})$. A candidate $G_k^{'}$ is sampled from a proposal distribution $q(G_k^{'}|G_k)$ and accepted with probability

$$\alpha = \min\left\{1, \frac{p(G_k^{'}|\mathbf{y}_{1:T},\mathbf{s}_{1:T})q(G_k|G_k^{'})}{p(G_k|\mathbf{y}_{1:T},\mathbf{s}_{1:T})q(G_k^{'}|G_k)}\right\} = \min\left\{1, \frac{p(G_k^{'}|\mathbf{y}_{1:T},\mathbf{s}_{1:T})p(G_k^{'})q(G_k|G_k^{'})}{p(G_k|\mathbf{y}_{1:T},\mathbf{s}_{1:T})p(G_k)q(G_k^{'}|G_k)}\right\}$$

This add/delete edge move proposal is rather accurate, although such accuracy comes at the price of a substantial computational burden. The full conditional posterior distribution of the state-specific SUR coefficient $\boldsymbol{\beta}_k$ is conjugate and defined as

$$p\left(\boldsymbol{\beta}_k|\Sigma_k,\mathbf{y}_{1:T},\mathbf{s}_{1:T}\right) \propto \mathcal{N}\left(M_k^*\left(\sum_{t\in\mathcal{T}_k} Z_t\Omega_k\mathbf{y}_t + M_k^{-1}\mathbf{m}_k\right), M_k^*\right) \tag{16}$$

with $M_k^* = \left(\sum_{t\in\mathcal{T}_k} Z_t\Omega_k Z_t^{'} + M_k^{-1}\right)^{-1}$, and $\Omega_k = \Sigma_k^{-1}$ is the inverse covariance matrix given the underlying graph structure, $G_k$. Finally, the conjugate Dirichlet prior for

the rows of the transition probability $\boldsymbol{\pi}_k = (\pi_{k1}, \ldots, \pi_{kK})'$ is updated as follows

$$p(\boldsymbol{\pi}_k|\mathbf{s}_{1:T}) \propto \mathcal{D}ir\left(c_{k1} + N_{k1}, \ldots, c_{kK} + N_{kK}\right) \tag{17}$$

with $N_{lk}$ is the empirical transition between the $l$th and the $k$th latent discrete states, i.e. $N_{lk} = \sum_{t=1}^{T} \xi_{lk,t}$ with $\xi_{lk,t} = \mathbb{I}_{\{k\}}(s_t)\,\mathbb{I}_{\{l\}}(s_{t-1})$.

## 3.3 States Identification via Eigenvector Centrality

The likelihood function in Eq.(13) remains unchanged with respect to any state permutation. Therefore, under an invariant prior specification, the posterior will also be invariant to any state permutation (see Frühwirth-Schnatter 2006 for a review of the label-switching and identification issues). We propose a state-identifying restriction based on a weighted eigenvector centrality measure. In its general form, the eigenvector centrality of the $i$th firm/stock, is a quantity proportional to the sum of the centralities of the neighbours of a vertex, so that a vertex may display a high centrality either by being connected to a lot of others or by being connected to others that themselves are highly central.

As in many other fields (e.g., biology, neuroscience, and operations research), where complex networks have been studied, it is possible to assign to each edge of the graph a weight proportional to the intensity of the connections among the various elements of the network. Appropriate metrics combining weighted and topological observables have been discussed in the literature (see Rubinov and Sporns 2010 for a review); as argued in Barrat, Barthlemy, Pastor-Satorras, and Vespignani (2004), network metrics based on weighted edges allow us to provide a better description of the hierarchies and organizational principles at the basis of the architecture of the networks. In the Gaussian graphical model literature, covariances, precisions, and the topological features of the graph, such as the paths between pairs of nodes, e.g., Jones and West (2005), have been routinely combined together with the goal of assessing the centrality of each node.

In our financial application, the information about the existence of a financial

linkage between pairs of stocks is encoded in the presence or absence of an edge between two nodes, while the strength of the linkage is measured by the covariance between pairs of stocks. In this respect, we construct a state-specific weighted graph defined as $\tilde{G}_k = (V_k, D_k, \tilde{\Sigma}_k)$, where $\tilde{\Sigma}_k$ is a real-valued symmetric matrix, called weight matrix, in which each entry is

$$\tilde{\sigma}_{ij,k} = \begin{cases} \sigma_{ij,k} & \text{if } (i,j) \in D_k \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

that is the weight $\tilde{\sigma}_{ij,k}$ assigned to each pair of nodes $\{i,j\} \in V_k \times V_k$ is equal to the corresponding covariance if they are connected in state $k \in K$. The relative centrality score for the $i$th firm in the weighted graph $\tilde{G}_k$, i.e. $\gamma_{i,k}$, can then be defined as:

$$\lambda_k \gamma_{i,k} = \sum_{j=1}^{n} \tilde{\sigma}_{ij,k} \gamma_{j,k} \tag{19}$$

where $\lambda_k$ is some constant. With a small re-arrangement, this measure can be re-written in a more compact form as the eigenvector equation $\tilde{\Sigma}_k \boldsymbol{\gamma}_k = \lambda_k \boldsymbol{\gamma}_k$ where $\boldsymbol{\gamma}_k = (\gamma_{1,k}, \ldots, \gamma_{n,k})'$. Since $\tilde{\Sigma}_k$ is real and symmetric, a unique solution is guaranteed to exist by the Perron-Frobenius theorem.[10] As a result, our average weighted eigenvector centrality of the graph $\tilde{G}_k$ is defined as

$$\overline{q}(\tilde{G}_k) = \frac{1}{n} \sum_{i=1}^{n} \gamma^*_{i,k} \tag{20}$$

where $\gamma^*_{i,k}$ is the $i$-th element of the eigenvector $\boldsymbol{\gamma}^*_k$ corresponding to the largest eigenvalue $\lambda^*_k$. The advantage of (20) is that it accounts not only for the number of connections of each node with the adjacent nodes, but also for its weight and for the weights of the indirect connections with other nodes of the graph. The notion of indirect connection between nodes can be more precise by means of the definitions of walks and paths between nodes of $\tilde{G}_k$.

[10] Since Bonacich (1972) proposed to use the eigenvector of the largest eigenvalue (maximal eigenvalue) as centrality measure, the eigenvector centrality is sometimes labelled Bonacich's centrality. See, e.g. Bollobás (1998), ch. 8, Th. 5 for a relationship between the maximal eigenvalue and the minimal and maximal degree of a graph.

**Definition 1.** *A walk* $p_{ij} = (i_0, e_1, \ldots, e_l, i_l)$ *between two vertices* $i$ *and* $j$ *of* $G_k$*, called endvertices, is identified by an alternating sequence of (not necessary different) vertices* $V_k(p_{ij}) = \{i_0, i_1, \ldots, i_l\}$ *and edges* $D_k(p_{ij}) = \{e_1, \ldots, e_l\} \subset D_k$*, with* $e_1 = (i_0, i_1)$*,* $e_l = \{i_{l-1}, i_l\}$*, and* $i_0 = i$ *and* $i_l = j$*. The number of edges* $|D_k(p_{ij})| = l$ *in a walk is called "walk length". A walk of length* $l$ *is called* $l$*-walk.*

**Definition 2.** *A path* $p_{ij}$ *between vertices* $i$ *and* $j$ *of* $G_k$ *is a walk with distinct elements in its vertex set. The shortest-path* $p^*_{ij}$ *between two vertices* $i$ *and* $j$ *is* $\min_l\{p_{ij} = (i_0, e_1, \ldots, i_l, e_l), l \geq 1\}$ *that is the path with the minimum length.*

See Bollobás (1998), ch. 1, for further details. Proposition 1 shows that our measure can be interpreted as a weighted sum over the walks where weights are inversely related to the length of each walk.

**Proposition 1.** *Let* $\tilde{G}_k = (V_k, D_k, \tilde{\Sigma}_k)$ *be a weighted undirected graph with vertex set* $V_k = \{1, \ldots, n\}$*, edge set* $D_k$ *and real-valued weight matrix* $\tilde{\Sigma}_k$*. Let us denote with* $\lambda_{1,k} \leq \lambda_{2,k} \leq \ldots \leq \lambda_{n,k}$ *the eigenvalues of* $\tilde{\Sigma}_k$*, with* $\boldsymbol{\gamma}_{j,k}$*,* $j = 1, \ldots, n$ *the associated eigenvectors and with* $\boldsymbol{\iota}$ *the* $n$*-dimensional unit vector. Assume the maximal eigenvalue* $\lambda_{n,k}$ *has multiplicity one, then the average eigenvector centrality measure* $\overline{q}(\tilde{G}_k)$ *satisfies the following*

$$\overline{q}(\tilde{G}_k) = \lim_{\beta \to 1/\lambda^-_{n,k}} \frac{1}{\kappa(\beta, k)} \sum_{l=1}^{\infty} \beta^{l-1} \tilde{\sigma}_{l,k} \tag{21}$$

*and*

$$\tilde{\sigma}_{l,k} = \sum_{i,j \in V_k} \sum_{p \in P^{(l)}_{ij,k}} \prod_{r=1}^{l} \sigma_{i_{r-1} i_r, k}$$

*is the sum of the walk weights over the set* $P^{(l)}_{ij,k}$ *of all possible* $l$*-walks between* $i$ *and*

*j, and,*

$$\kappa(\beta, k) = \frac{1}{n} \sum_{i=1}^{n} \frac{\beta \lambda_{i,k}}{1 - \beta \lambda_{i,k}}$$

*is a normalizing factor with* $|\beta \lambda_{i,k}| < 1$.

*Proof* See Appendix A.

Other centrality measures have been proposed in the literature. These include average node degree, closeness, and betweenness. However, these measures may fail in identifying alternative regimes of systemic risk. Degree centrality gives a simple count of the number of connections any given firm or asset maintains, without effectively discriminating the relative importance of these connections with respect to the whole network. However, linkages across firms are not all alike: for instance, firms in large sectors such as "industrials" are likely highly connected to other firms through obvious supply relationships; this implies that, e.g., a demand shock to Fedex could be more systemically important than a liquidity shock to JP Morgan, which is not necessarily the case. Analogously, closeness and betweeness are measures of centrality based on shortest paths between the node and all other nodes in the graph. These measures implicitly assume simplistic and pre-determined paths and may be severely inappropriate when applied to economic shocks. Only shocks with known destination follow the shortest possible distance (e.g., when a firm purposely alters a relationship with another firm to affect the stance of the latter). Instead, economic shocks are in general unlikely to be restricted to follow specific paths but are also likely to have feedback effects. For instance, a liquidity shock hitting a specific debtor firm could affect its ability to pay back a loan to a bank, that could prevent the bank to extend a credit line to another firm, that in turns will no longer afford to pay for its supply debts; eventually, such a feedback could flow back to hit the original firm if the first firm originating the shock and the latter were to be in a trading relationship.

Our measure addresses this issue that the existence of an edge or a shortest path between two nodes do not necessarily measure the level of connectedness between two nodes. In this respect, our centrality measure shares some features with the communi-

cability measures for complex networks proposed in Estrada and Hatano (2008, 2009) and more generally with other global connectivity measures (e.g., Qi, Fuller, Wu, Wu, and Zhang 2012, Qi, Fuller, Luo, and Zhang 2015 and Han, Escolano, Hancock, and Wilson 2012) admitting decompositions in walks, spanning trees, and circuits. More specifically, the following decomposition in shortest paths and walks holds true for our weighted eigenvector centrality measure.

**Corollary 1.** *Let* $\tilde{\sigma}(p) = \prod_{r=1}^{l} \sigma_{i_{r-1}i_r,k}$ *be the weight products over the edges of walk* $p_{ij} = (i_0, i_2, \ldots, i_l)$*, then the weighted eigenvector centrality measure* $\overline{q}(\tilde{G}_k)$ *can be written as*

$$\overline{q}(\tilde{G}_k) = \lim_{\beta \to 1/\lambda_{n,k}^-} \frac{1}{\kappa(\beta, k)} \sum_{i,j \in V_k} \left( \sum_{p \in P_{ij,k}^*} \beta^{l_{ij}^* - 1} \tilde{\sigma}(p) + \sum_{l > l_{ij}^*} \beta^{l-1} \sum_{p \in P_{ij,k}^{(l)}} \tilde{\sigma}(p) \right) \tag{22}$$

*where* $P_{ij,k}^*$ *and* $l_{ij}^*$ *denotes the set of the shortest paths between nodes* $i$ *and* $j$ *of the graph* $\tilde{G}_k$ *and their length, respectively, and* $\in P_{ij,k}^{(l)}$ *denotes the set of walks from* $i$ *to* $j$ *for a given state* $k = 1, \ldots, K$*.*

*Proof* See Appendix A.

The first term in Eq.(22) reflects the connectivity due to the shortest paths and degree distribution, while the second terms reflects the connectivity or influence between nodes at a global level, reflecting losses spreading into the financial system forward and backward several times from a source to a destination. Thus, the measure we propose provides a better representation of more complex structures such as scale-free or small-world networks. This aspect is also reflected by a different decomposition of our measure into intra-cluster and inter-cluster communicability terms.

**Corollary 2.** *The weighted eigenvector centrality measure* $\overline{q}(\tilde{G}_k)$ *can be written as*

$$\overline{q}(\tilde{G}_k) = \lim_{\beta \to 1/\lambda_{n,k}^-} \frac{1}{\beta \kappa(\beta, k)} \sum_{i \in V_k} \frac{\beta \lambda_{i,k}}{1 - \beta \lambda_{i,k}} \left( \varphi_{i,k}^{(1)} - 2\varphi_{i,k}^{(2)} \right) \tag{23}$$

*where* $\varphi_{i,k}^{(1)} = (\boldsymbol{\gamma}_{i,k}^{+\prime} \boldsymbol{\iota})^2 + (\boldsymbol{\gamma}_{i,k}^{-} \boldsymbol{\iota})^2$ *and* $\varphi_{i,k}^{(2)} = \boldsymbol{\gamma}_{i,k}^{+\prime} \boldsymbol{\iota} \boldsymbol{\gamma}_{i,k}^{-\prime} \boldsymbol{\iota}$ *with* $\boldsymbol{\gamma}_{i,k}^{+}$ *and* $\boldsymbol{\gamma}_{i,k}^{-}$ *the element-wise*

*positive and negative parts of* $\boldsymbol{\gamma}_{i,k}$*, respectively.*

*Proof* See Appendix A.

Intuitively if the $i$ and $j$ entries of a eigenvector have the same sign, then the corresponding nodes react in a similar way to a shock propagating through the network (see, e.g. Estrada and Hatano 2008). Thus, the nodes can be partitioned into groups following the sign of their overall contribution $(\varphi_{i,k}^{(1)} - 2\varphi_{i,k}^{(2)})$ to the centrality measure. The decomposition shows that our measure also naturally accounts for community structures (see Fortunato 2010 for a graph-theoretic definition of community).

At this point, we have completely developed the apparatus necessary to our state identification strategy based on network statistics. Regimes identification is now based on the restriction

$$\overline{q}(\tilde{G}_1) < \ldots < \overline{q}(\tilde{G}_K),$$

This constraint allows us to directly "separate" regimes according to the density of the network. In this respect, one of the major advantages of this identification scheme is that we can give a clear economic interpretation to the regimes: the first regime is associated with the lowest level of systemic risk and hence the lowest average of centrality scores across firms, the second regime with the next lowest average incidence of systemic risk, and so on, with the last regime associated with the strongest incidence of systemic risk. Appendix B shows the efficiency of the MCMC posterior approximation procedure based on simulated datasets.

# 4 Empirical Analysis

Our application focusses on all the constituents stocks of the S&P100 index for which we have at least fifteen years of continuous trading days as of the end of our sample (this makes our cross-section equal to $n = 83$). The sample period is May 1996 - October 2014. The constituents of the S&P 100 represent about 63% of the market capitalization of the S&P 500 and about a half of the total market capitalization of

the U.S. equity markets as of January 2017. These stocks tend to be the largest and most established companies in the U.S., which mitigates liquidity concerns.

We analyze three popular linear asset pricing models starting from the plain vanilla CAPM, then extended to the three-factor model proposed by Fama and French (1993) to include both size and value factor-mimicking portfolios, to conclude with a recently offered implementation of the Merton (1973) intertemporal CAPM in which the aggregate dividend yield and fixed income default and term spread are considered as state variables in addition to the excess returns on the market portfolio. The default spread is computed as the difference between the yields of long-term corporate Baa bonds and long-term government bonds and should reflect a risk premium for the aggregate risk of firm's default on their debt obligations. The term spread is measured as the difference between the yields on 10- and 1-year government bonds, and reflects the slope of the risk-free yield curve, a well-know business cycle leading indicator.[11] For ease of exposition, we report the results for the three-factor Fama-French model and the I-CAPM. In fact, the results for the CAPM are largely overlapping with the I-CAPM.[12]

## 4.1 Estimates of Latent States and Parameters

A priori, we assume that the latent states are persistent. This is based on the conventional wisdom that posits that systemic risk is not a quickly mean-reverting process (see, e.g., Forbes and Rigobon 2002, Billio et al. 2012 and Diebold and Yilmaz 2014). In our application, we set the hyper-parameters to be rather uninformative: $m_k = 0$ and $M_k = 1000 I_n$ for each $k = 1, \ldots, K$, i.e., the prior structure is assumed not to differ across low vs high systemic risk states.

[11] Data on corporate bonds and Treasuries are from the Fred II database of the Federal Reserve Bank of St.Louis. Following Campbell (1996) in the I-CAPM implementation, we use as the state variables (factors) the innovations estimated from a first order Vector Auto-Regressive VAR(1) process. Thus, for each collection of the CRSP aggregate value-weighted market portfolio and the candidate set of risk factors $\mathbf{h}_t$, we estimate $\mathbf{h}_t = B_0 + B_1 \mathbf{h}_{t-1} + \boldsymbol{\eta}_t$ for $t = 1, \ldots, T$. Following Petkova (2006), the innovations $\boldsymbol{\eta}_t$ are orthogonalized relative to the excess return on the aggregate wealth and scaled to have the same variance.

[12] Simple CAPM results are available from the authors upon request. Note that the reduced-form nature of our framework necessarily makes all empirical results sample-specific.

The prior for the Hyper-Inverse Wishart distribution is also set to be fairly vague, albeit proper, by selecting $a_k = 3$ and $A_k = 0.0001 I_n$. Finally, the marginal prior for the graph space is a Bernoulli distribution with $\psi = 2/(n-1)$ which would provide a prior mode at $n$ edges. In order to further reduce the sensitivity of posterior estimates to the prior specification, we use the initial 20% of the draws as burn-in sample.

In Appendix C we formally test for the number of regimes; evidence from Bayes factors are clearly in favour of a specification with two-regimes. Figure 1 and 2 show the filtered probability of being in a state of high systemic risk across different regression specifications as well as the transition probabilities of the latent states, respectively. The gray area represents the systemic risk probability, while the red solid line shows the NBER recession indicator for the period following the peak of the recession to the trough.

[Insert Figure 1 and 2 about here]

Figure 1 shows that high network connectivity has characterized the period 2001-2002 (i.e., the dot.com bubble, the market fears that followed the September 2001 terror attacks, the Enron and Worldcom financial scandals, and the unfolding of the events leading to the second Iraqi war), and the great financial crisis of 2008-2009. Although there is an obvious mis-matching between the ex-post identification of the high network connectivity regime and the NBER business cycle indicator over the period 1998-2002, the NBER recession and high systemic risk tend to overlap across the recent great financial crisis. As a whole, our filtered probabilities line up fairly well with well-known periods of increasing turmoil in financial markets.

One comment is in order; in this paper we are not focussing on financial crises per sé. The fact that few events at the beginning of 2000 made the economic system more fragile and therefore could have been reflected in cross-sectional connectedness is highly possible even though they have not, or not always, triggered full-blown financial crisis episodes. Interestingly, we show below that also outside episodes well known to represent financial crises, it turns out that financial firms (besides a few specific sectors) always come to occupy a central positioning in the network because

of their high connectivity. In this respect, the outcome that financial firms are more relevant when aggregate connectivity is high is endogenously generated and not driven by the selection of stocks in the financial industry as test assets.

Figure 2 shows standard confidence box plots for the posterior transition probabilities across different regression specifications. The first (last) three columns represent the probability of staying in a state of low (high) systemic risk. Under all regression specifications, states of high systemic risk are rather persistent, with an approximate median probability of $\pi_{hh} = 0.9$. The probability of staying in a regime of high network connectdness is slightly lower than the corresponding probability of a lack of systemic risk to persist, which is compatible with the empirical evidence that turbulent periods of crisis and contagion tend to be less persistent than "normal" times.

Figure 3 and 4 plot the posterior of the between-state differences of conditional intercepts, i.e., the so-called Jensen's alphas, and betas from the Fama-French three-factor model and an I-CAPM implementation, respectively. For each stock in our sample, we report the box-plot of the posterior distributions of the parameters. For ease of exposition, we cluster the ticks on the horizontal axis according to an industry classification, in the sense that posterior of the different coefficients for each of the stocks are grouped around a tick that indicates the industry they belong to. The top-left panel shows the difference in the regime-specific Jensen's alphas for the three-factor model. Interestingly, we estimate substantial changes in the abnormal average pricing error across different regimes of network connectivity. Such differences, especially those derived from the I-CAPM, are predominantly positive, i.e., alphas tend to be higher in calm, low systemic risk environments, when as a result, factor pricing models seem to systematically under-estimate average excess returns. Therefore it is when the meta-factor represented by systemic risk variations is added into the picture that the factor models are providing the best performance, in terms of lower Jensen's alphas. Moreover, the exposure to market risk (top-right panel) increases quite dramatically in states of high systemic risk and this tends to occur across most stocks, even though the effect is especially visible for stocks from the financial sector,

when the increase in mean posterior beta is in the order of almost 0.3. A similar increasing exposure to market risk, is also estimated for some stocks in the Materials and Tech industries. Similarly, the finance sector seems to be more exposed to the value risk factor (bottom-right panel) when aggregate connectedness is estimated to be higher. The Industrial and Materials sectors also show an increasing exposure to value mimicking factor portfolio when systemic risk is higher.

[Insert Figures 3 and 4 about here]

Figure 4 shows similar results for the I-CAPM specification. The market betas (top-right panel) behave in a way that is hard to tell apart from the Fama-French three-factor regression specification. Unlike the three-factor model, the Energy sector shows now that the impact of market risk tend to be lower when network connectivity is high. The bottom-left panel shows the change of betas on default risk. On average, the exposure to default risk is higher when systemic risk is high, although for a large fraction of the sample such negative relationship is hardly significant, in the sense that posterior confidence regions often include a zero difference in betas. The bottom-right panel shows that stock exposures to dividend yield risk tend to display instead a counter-cyclical effect for both the Energy and Financial sectors, in the sense that betas are higher during periods of crisis and high systemic risk. Also Tech firms tend to react more to a surprise in the dividend yield in periods of high systemic risk.

## 4.2 Network Estimates

Under the MCMC estimation scheme outlined in section 3, it is possible to define the posterior distribution of the graph and covariances $p(\Sigma_k, G_k|\mathbf{y}_{1:T})$ and to assess the statistical properties of the firm-specific network contribution using equation (20). Denote as $\overline{q}_k$ our weighted eigenvector centrality measure. Its posterior distribution

can be approximated as;

$$p_J(\overline{q}_k|\mathbf{y}_{1:T}) = \frac{1}{J}\sum_{j=1}^{J}\delta_{\overline{q}_k^j}(\overline{q}_k) \tag{24}$$

where $\delta(\cdot)$ is a Dirac's delta function, $q_k^j$ the average weighted centrality measure which takes as input the state-dependent graph $G_k^{(j)}$ and covariance $\Sigma_k^{(j)}$ sampled from the posterior distribution, and $J$ is the number of Gibbs iterations. We explicitly account for the posterior uncertainty associated with the weighted graph $\tilde{G}_k$ by using the integrated measure

$$\int\int_{\mathcal{M}(G_k)\times\mathcal{G}} \overline{q}(\tilde{G}_k)p(\Sigma_k, G_k|\mathbf{y}_{1:T})d\Sigma_k dG_k \approx \int_Q \overline{q}_k\, p_J(\overline{q}_k|\mathbf{y}_{1:T})d\overline{q}_k \tag{25}$$

which is the empirical average of the sequence of measures $\overline{q}_k^j$, $j = 1, \ldots, J$, associated with the MCMC draws. We first display the network implied by the posterior median graphs computed according to equation (25). Figures 5-6 show the results obtained from the Fama-French three-factor model and the I-CAPM implementation, respectively. The size and the color of the nodes are proportional to their relevance in the network measured by weighted eigenvector centrality. The darker (bigger) the color (size) of the node, the higher its marginal contribution to aggregate systemic risk.

[Insert Figure 5 and 6 about here]

A few comments are in order. First, the strength of cross-firm connectedness differs substantially across different specifications of the factor models. The network tends to be more sparse for the I-CAPM where more state variables are included. This is possibly due to the residual nature of the estimated network. In this respect, by including a significant risk factor the ability of our model to clear the network structure from spurious linkages increases. As shown in Figure 4 the betas on dividend-yield for the I-CAPM tend to be highly significant in the cross-section of firms, which implies that, unlike the Fama-French specification, the network from the I-CAPM is conditional on the aggregate market under- or over-valuation. Second, financial

firms are increasingly pivotal to the network during periods of high systemic risk. This is consistent with the conventional wisdom that posits that the financial sector is central in the transmission mechanism of exogenous shocks during crises. Third, firms within the Energy sector show instead the highest degree of network centrality in more tranquil periods.

The role of energy firms is not entirely unexpected. Historically, a number of aggregate shocks to the US economy – that was otherwise close to a full-employment path – have come from energy shocks and most or all such shocks did come as a genuine surprise (as triggered by international conflicts or by OPEC's decisions), it is not completely surprising that we find that in good times, periods of low systemic shocks, energy firms may occupy a more central position in the estimated network.

Figure 1 and Figures 5-6 combined, confirm that during periods of market turmoil, the systemic importance of the financial sector substantially increases. Finally, the marginal importance of each firm in the network is confirmed even explicitly conditioning on size and book-to-market beta exposures as sources of risk. In this respect, the key role of the Financial (Energy) sector when systemic risk is high (low) is confirmed even when size and value exposures are netted out.

We now shift our attention to the contribution of individual stocks/firms to aggregate systemic risk. Figure 7 shows the top 20 stocks ranked according to the posterior median of the weighted eigenvector centrality measure in Eq.(19). The red (blue) line with circle (square) marks shows the centrality measure across companies when aggregate network connectedness is low (high).

[Insert Figure 7 about here]

The left panel shows the results for the Fama-French three-factor model. In a state of low aggregate network connectivity, energy stocks such as Exxon Mobil (XOM) and Schlumberger (SLB), tend to carry the highest weight in the system. Moreover, financials stocks rank at the top positions in terms of their weight in our estimates of systemic risk in a state of high aggregate connectedness. For instance, Bank of

America (BAC) has double the effect of Exxon Mobil (XOM) and for times the weight of ConocoPhillips (COP) on system risk. The right panel of Figure 7 extends these results to the I-CAPM model. First, by conditioning on macroeconomic-oriented factors, the relative contribution of energy companies becomes now lower. Companies such as Anadarko Ptl. (APC), ConocoPhillips (COP), Occidental Ptl. (OXY), Apache (APA), and Schlumberger (SLB) show now a much lower centrality in the posterior estimated network. This suggests that much of the effect of energy companies in the three-factor model may be due to the fact that traded factors cannot accurately capture and represent the state of the macroeconomy, which in its turn appears to be inherently related to energy shocks. Analogously to the Fama-French model, however, when aggregate systemic risk is higher (blue line), the weight of financial institutions in the network tend to dominate all other industries.

These measures of firm-specific eigenvector centrality can be generalized at the industry level by averaging $\boldsymbol{\gamma}^*_{i,k}$ within a certain industry. For instance, for the financial sector can be approximated as

$$\boldsymbol{\gamma}_{fin,k} = \frac{1}{n_{fin,k}} \sum_{i \in D_{fin,k}} \boldsymbol{\gamma}^*_{i,k} \tag{26}$$

with $n_{fin} = |D_{fin,k}|$, and $D_{fin,k} \subset D_k$ the set of nodes associated to firms classified in industry groups according to the Global Industry Classification Standard (GICS), developed by MSCI. We use the first layer of classification which consists of 11 sectors into which S&P has categorized all major public companies. Thus, none of the sectors is composed by just 2-3 companies (let alone one). Such wide classification coupled with the fact that the S&P 100 represent about 63% of the market capitalization of the S&P 500 likely make our industry-based results robust, at least in their qualitative implications, with respect to more broad indexes, e.g. S&P500. The industry-level centrality measures are obtained by taking the median of firm-specific measures averaged out within industries. Figure 8 shows the results; left (right) panel shows the results for the state of low (high) aggregate connectivity.

[Insert Figure 8 about here]

Both the financial and the energy sectors tend to dominate across regimes. When aggregate connectedness is high, the importance of industries such as utilities, telecommunications, health care, consumer staples, and discretionary consumption goods are almost negligible. This is due to a substantial concentration of the network around few firms that belong to either the financial and the energy industries. As a whole, although the visual representation of the network graphs in Figures 5-6 may change across factor pricing models, key features such as the ranking of firms and industries in terms of their systemic importance are robust across factor pricing model specifications. A possible explanation is based on the dominance of the market risk factor, which is included as independent variable in all of our regression specifications. In this respect, while additional factors might help to refine the analysis of the graph network, the key properties of cross-firm connectedness depend on the fact that a benchmark market portfolio is included in the set of risk factors.[13]

## 4.3 Market Value, Financial Losses, and Network Centrality

The network centrality of a firm/industry is possibly linked to its relative market value. For instance, the relative equity weight of the financial sector has dropped from 20% in 2006 to less than 10% by the end of the financial crisis of 2008-2009, when network connectedness had been high (see Figure 1). This implies the existence of an inverse relationship between the centrality of the financial sector and its corresponding market value. The opposite is true for the energy sector: the relative market value of the energy sector has increased through our sample and has tended to becomer higher during periods characterized by high aggregate network connectivity. The same positive relationship can be seen to apply, although to a weaker extent, to telecommunication service stocks, while listed companies belonging to industrial and material sectors do not display a clear mapping between their relative importance and aggregate systemic risk. Also, the relative market value of the technology industry did

[13]Notice that the results that we have obtained in the application are necessarily sample-specific. It is well possible that by ending our data sample before the Great Financial Crisis, the overall picture concerning which firms are leading because of their network centrality may change. Once more, this has to do with the fact that our framework is based on a reduced-form multi-factor pricing model which does not allow to perform structural analysis.

considerably increase during the late 1990s to then deflate back to the original values beginning in 2000, in correspondence to the burst of what has been then defined to be a dot.com bubble.

Because these empirical facts are rather stark, we also proceed to test for the existence of a significant relationship between firm-level network centrality and market values across regimes. We do so by estimating a set of univariate cross-sectional regressions where the dependent variable is the centrality measure for each firm in regime $k$, and the explanatory variable is the corresponding market value averaged across the periods identified by regime $k$. We compute such regression for each factor pricing model and the two different regimes isolated early on. For each regression, we control for industry heterogeneity by including a fixed effect that identifies the industry of where each firm belongs. We report the regression coefficient, its t-statistic, and the adjusted $R^2$, together with a rank-correlation coefficient as in Kendall (1938). We first rank stocks according to their centrality within the network and then we rank them according to their average market value across regimes. Kendall's $\tau$ coefficient measures the correspondence between the two rankings. Table 1 shows the empirical results.

[Insert Table 1 about here]

We find evidence that systemic risk and market value are not decisively correlated. Top panel shows the results for our weighted centrality measure. The slope coefficient is low in magnitude and not statistically significant across regimes. The t-statistics are below the 5% significance threshold, and the adjusted $R^2$ is below 3% across models and regimes.

One may argue that this finding does not square well with other evidence in the literature especially in the banking and insurance sectors. Indeed, except for the peak of the financial crisis episodes, the empirical evidence provided so far in the literature tend to be a positive association between the contribution to systemic risk and market capitalization or the asset value of a firm. However, on the one hand, it should be clear that our concern was not to over-throw a result in the empirical finance literature, but to reject a concern that our empirical network analysis may just represent an

involute and complex way to talk about firm and/or sector value-weighted size, to the borders of triviality. Our results in Table 1 show that this is not the case: there is only weak correlation between market value and network connectedness, although such correlation is indeed positive. On the other hand, clearly the largest financial institutions caused major systemic effects (even though the largest among the US banks, Citigroup, in fact has been little discussed and mentioned in accounts of the crisis, even though its role in the US financial network and implied systemic risk begs no doubts, we trust). Here the re-solution of the apparent puzzle lies in two facts: first, our paper does not concern stock market crises, it is more generally about the estimation of systemic weakness through network connectedness; as a result it is possible that accounts from memorable crisis episodes may concern just a subset of the relevant dynamic linkages that we have isolated. Second, although the financial crisis looms large in our memories and was actually characterized by a major role by the large institutions, there have been other, earlier crises triggered by smaller companies (also within the financial sector, think of the case of Long-Term Capital Management default) and sectors characterized by smaller companies (think of the dot-com crisis).

One further implication for any systemic risk measure is its ability to accurately correlate with potential losses experienced by firms. To this end, we test the null that those firms more exposed to systemic risk are those that tend to have higher losses. For each model and regime we regress the average maximum percentage financial loss (AM%L henceforth) onto the network centrality measure for $i = 1, \ldots, N$ firms.[14] Again, we control for industry heterogeneity by inclusing an industry fixed effect. The results are reported in Table 2.

[Insert Table 2 about here]

We find that companies more exposed to the overall systemic risk of the network, i.e.,

[14]Suppose that a regime of high systemic risk lasts from $t$ to $t + h$. The maximum percentage loss for a firm is defined to be the maximum difference between the market capitalization of an institution at time $t$ and $t + h$ dividend by its market capitalization at time $t$. The average measure is computed by averaging out such maximum percentage loss across those periods identified by the hidden state $s_t$.

those with higher weighted eigenvector centrality, are more likely to suffer significant losses and hence a declining relative market value, when aggregate systemic risk is larger. In this respect, our centrality measure is similar to the marginal expected shortfall (MES) measure originally proposed by Acharya, Pedersen, Phillippon, and Richardson (2017), which tracks the sensitivity of returns on stock $i$ to a system-wide extreme event, thereby providing a market-based measure of fragility of a firm. Table 2 shows indeed that firms that are more contemporaneously interconnected with the rest of the market are also those that experience major losses in terms of market valuation in periods where such interconnections matter because systemic risk is high. The cross-sectional regression coefficient is indeed significant at standard confidence levels and the adjusted $R^2$ is around 12% across models. However, such positive correlation between network centrality and market losses is marginally less significant when aggregate connectedness decreases, i.e., in states of low systemic risk.

Table 2 also reports a Kendall's rank-correlation coefficient obtained by ranking firms from 1 to $N$ according to their centrality first and then according to the opposite of their cumulative returns (losses) suffered across regimes. The rank correlation results confirm that there is a significant relationship between network centrality and value losses across firms, especially during periods of high aggregate systemic risk, i.e., firms more exposed to systemic risks will face larger losses on average. This is consistent with previous empirical evidence in Billio et al. (2012) and Diebold and Yilmaz (2014), and the theoretical framework of Acemoglu et al. (2012).

# 5 Conclusions

In the aftermath of the great financial crisis, one of the main questions for economists and market participants has concerned the extent to which the economy is robust to unexpected shocks. In the language of network analysis, this translates into a desire to understand the nature and density of cross-firm connectivity. The concern of researchers and policy-makers alike is whether and when it may be possible for a shock originating in a corner of a complex web of economic and financial market con-

nections to initiate aggregate, systemic shocks with a large, widespread impact. We address this question by developing a novel Markov Switching Graphical Seemingly Unrelated model, which allows us to jointly estimate standard SUR-type relationships along with firms' or asset prices' connectedness from the error terms of linear multi-factor pricing model specifications. By conditioning on different sources of systematic risk, we implicitly recognize that systematic and systemic risk might be conditionally independent but not mutually exclusive.

Methodologically, we develop a Markov Chain Monte Carlo (MCMC) scheme which allows to sample the posterior estimates of all the variables of interests. The label-switching identification problem is solved using the graph-theoretic properties of the state-specific conditional dependence structures of the regression residuals. In this respect, we propose a new weighted eigenvector centrality measure, which accounts not only for the number of adjacent nodes, but also for the weights of the edges and for the number of indirect connections between nodes. More generally, our new measure implies that the existence of a financial linkage between two firms is encoded in the presence or absence of an edge, while the strength of the linkage is measured by their covariance.

# References


Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi. 2012. The network origins of aggregate fluctuations. *Econometrica* 80:1977–2016.

Acharya, V., L. Pedersen, T. Phillippon, and M. Richardson. 2017. Measuring Systemic Risk. *Review of Financial Studies* 30:2–47.

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

Ang, A., and G. Bekaert. 2002. International asset allocation with regime shifts. *Review of Financial Studies* 15:1137–1187.

Barigozzi, M., and C. Brownlees. 2014. NETS: Network Estimation for Time Series. *Unpublished Working Paper* .

Barrat, A., M. Barthlemy, R. Pastor-Satorras, and A. Vespignani. 2004. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America* 101:3747–3752.

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

Bollobás, B. 1998. *Modern Graph Theory*. Springer.

Bollobás, B. 2001. *Random Graphs*. Cambridge University Press.

Bonacich, P. 1972. Factoring and weighting approaches to clique identification. *Journal of Mathematical Sociology* 2:113–120.

Bonacich, P. 2007. Some unique properties of eigenvector centrality. *Social Networks* 29:555–564.

Brownlees, C., E. Nualart, and Y. Sun. 2014. Realized Networks. *Unpublished Working Paper* .

Campbell, J. 1996. Understanding Risk and Return. *Journal of Political Economy* 104:298–345.

Carter, C., and R. Kohn. 1994. On Gibbs sampling for state-space models. *Biometrika* 81:541–553.

Carvalho, C., and M. West. 2007. Dynamic Matrix-Variate Graphical Models. *Bayesian Analysis* 2:69–98.

Carvalho, C., M. West, and H. Massam. 2007. Simulation of Hyper-Inverse Wishart Distributions in Graphical Models. *Biometrika* 94:647–659.

Casella, G., and C. P. Robert. 2004. *Monte Carlo Statistical Methods*. New York: Springer Verlag.

Chib, S. 1995. Marginal Likelihood from the Gibbs Output. *Journal of the American Statistical Association* 90:1313–1321.

Cochrane, J. 2001. *Asset Pricing*. Princeton, NJ: Princeton University Press.

Cont, R., and J.-P. Bouchaud. 2000. Herd Behavior and Aggregate Fluctuations in Financial Markets. *Macroeconomic dynamics* 4:170–196.

Corsetti, G., M. Pericoli, and M. Sbracia. 2005. Some Contagion, some Interdependence: More Pitfalls in Tests of Financial Contagion. *Journal of International Money and Finance* 24:1177 – 1199.

Corsetti, G., M. Pericoli, and M. Sbracia. 2011. *Financial Contagion: The Viral Threat to the Wealth of Nations*, chap. Correlation Analysis of Financial Contagion, pp. 11–20. John Wiley & Sons.

Dawid, A., and S. Lauritzen. 1993. Hyper-Markov Laws in the Statistical Analysis of Decomposable Graphical Models. *The Annals of Statistics* 21:1272–1317.

Dempster, A. 1972. Covariance Selection. *Biometrics* 28:157–175.

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

Diebold, F., and K. Yilmaz. 2015. *Financial and Macroeconomic Connectedness: A Network Approach to Measurement and Monitoring*. Oxford University Press.

Erdös, P., and A. Rényi. 1959. On the Random Graphs. *Publ. Math. Debrecen* 6:290–297.

Estrada, E., and N. Hatano. 2008. Communicability in complex networks. *Physical Review E* 77:036111–12.

Estrada, E., and N. Hatano. 2009. Communicability graph and community structures in complex networks. *Applied Mathematics and Computation* 214:500–511.

Fama, E., and K. French. 1993. Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33:3–56.

Forbes, K., and R. Rigobon. 2000. *Measuring Contagion: Conceptual and Empirical Issues*. International financial contagion ed. Kluwer Academic Publisher.

Forbes, K., and R. Rigobon. 2002. No Contagion, Only Interdependence: Measuring Stock Market Comovements. *Journal of Finance* 57:2223–2261.

Fortunato, S. 2010. Community detection in graphs. *Physics Reports* 486:75 – 174.

Frühwirth-Schnatter, S. 1994. Data Augmentation and Dynamic Linear Models. *Journal of Time Series Analysis* 15:183–202.

Frühwirth-Schnatter, S. 2006. *Finite Mixture and Markov Switching Models*. Berlin: Springer-Verlag.

Giudici, P., and P. Green. 1999. Decomposable Graphical Gaussian Model Determination. *Biometrika* 86:785–801.

Guidolin, M., and A. Timmermann. 2008. International asset allocation under regime switching, skew and kurtosis preferences. *Review of Financial Studies* 21:889–935.

Hamilton, J. 1994. *Time series analysis*. Princeton University Press.

Hammersley, J., and P. Clifford. 1971. Markov Fields on Finite Graphs and Lattices. *Unpublished Manuscript, Oxford University* .

Han, L., F. Escolano, E. R. Hancock, and R. C. Wilson. 2012. Graph characterizations from von Neumann entropy. *Pattern Recognition Letters* 33:1958 – 1967.

Hautsch, N., J. Schaumburg, and M. Schienle. 2015. Financial network systemic risk contributions. *Review of Finance* 19:685–738.

Jones, B., C. Carvalho, A. Dobra, C. Hans, C. Carter, and M. West. 2005. Experiments in Stochastic Computation for High-Dimensional Graphical Models. *Statistical Science* 20:388–400.

Jones, B., and M. West. 2005. Covariance decomposition in undirected Gaussian graphical models. *Biometrika* 92:779–786.

Kass, R., and A. Raftery. 1995. Bayes Factors. *Journal of the American Statistical Association* 430:773–795.

Kendall, M. 1938. A New Measure of Rank Correlation. *Biometrika* 30:81–89.

Lauritzen, S. 1996. *Graphical Models*. Claredon Press, Oxford.

Liew, J., and M. Vassalou. 2000. Can book-to-market, size and momemtum be risk factors that predict economic growth? *Journal of Financial Economics* 57:221–245.

Markose, S. M. 2005. Computability and Evolutionary Complexity: Markets as Complex Adaptive Systems (CAS). *The Economic Journal* 115:F159–F192.

Merton, R. 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41:867–887.

Petkova, R. 2006. Do the Fama-French Factors Proxy for Innovations in Predictive Variables? *Journal of Finance* 61:581–612.

Qi, X., E. Fuller, R. Luo, and C.-Q. Zhang. 2015. A novel centrality method for weighted networks based on the Kirchhoff polynomial. *Pattern Recognition Letters* 58:51–60.

Qi, X., E. Fuller, Q. Wu, Y. Wu, and C.-Q. Zhang. 2012. Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences* 194:240 – 253.

Roberts, G., and S. Sahu. 1997. Updating schemes, covariance structure, blocking and parametrisation for the Gibbs sampler. *Journal of the Royal Statistical Society* 59:291 – 318.

Rodriguez, A., A. Lenkoski, and A. Dobra. 2011. Sparse covariance estimation in heterogeneous samples. *Electronic Journal of Statistics* 5:981.

Rubinov, M., and O. Sporns. 2010. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* 52:1059 – 1069.

Timmermann, A., D. Blake, I. Tonks, and A. Rossi. 2014. Network Centrality and Fund Performance. *CFR Working Papers N 15-16, University of Cologne, Centre for Financial Research* .

Vassalou, M. 2003. News Related to future GDP growth as a risk factor in equity returns. *Journal of Financial Economics* 68:47–73.

Wang, H. 2010. Sparse Seemingly Unrelated Regression Modelling: Applications in Finance and Econometrics. *Computational Statistics and Data Analysis* 54:2866–2877.

Wang, H., C. Reeson, and C. M. Carvalho. 2011. Dynamic Financial Index Models: Modeling Conditional Dependencies via Graphs. *Bayesian Analysis* 6:639–664.

Wang, H., and M. West. 2009. Bayesian Analysis of Matrix Normal Graphical Models. *Biometrika* 96:821–834.

Whittaker, J. 1990. *Graphical Models in Applied Multivariate Statistics*. Chichester, UK: John Wiley and Sons.

# Appendix

## A Proofs of the results of Section 3

### A.1 Proof of the result in Eq. 14

The complete likelihood of the data is then defined as

$$p\left(\mathbf{y}_{1:T}, \mathbf{s}_{1:T}|\boldsymbol{\theta}, G\right) = \prod_{t=1}^{T} p\left(\mathbf{y}_t|s_t, \boldsymbol{\theta}, G\right) p\left(s_t|s_{t-1}, \boldsymbol{\theta}\right) p\left(s_0\right) \tag{A.27}$$
$$= \prod_{t=1}^{T} \left(2\pi\right)^{-n/2} \left|\Sigma(s_t)\right|^{-1/2} \exp\left(-\frac{1}{2}\left(\mathbf{y}_t - Z_t'\boldsymbol{\beta}(s_t)\right)' \Sigma_t^{-1} \left(\mathbf{y}_t - Z_t'\boldsymbol{\beta}(s_t)\right)\right) \prod_{k,l=1}^{K} \pi_{lk}^{\xi_{lk,t}} \pi_0,$$

with $\xi_{lk,t} = \mathbb{I}_{\{k\}}\left(s_t\right) \mathbb{I}_{\{l\}}\left(s_{t-1}\right)$. Thus the full conditional of $\Sigma_k$ given $\mathbf{y}_{1:T}$, $\boldsymbol{\theta}$, $\mathbf{s}_{1:T}$, is

$$p\left(\Sigma_k|\mathbf{y}_{1:T}, \boldsymbol{\theta}, \mathbf{s}_{1:T}, \boldsymbol{\beta}_k\right) \propto \prod_{t=1}^{T} \left(2\pi\right)^{-n/2} \left|\Sigma(s_t)\right|^{-1/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\Sigma(s_t)^{-1}\mathbf{e}_t\mathbf{e}_t'\right)\right\} p(\Sigma_k)$$
$$\propto \prod_{t\in\mathcal{T}_k} \left|\Sigma_k\right|^{-1/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\Sigma_k^{-1}\mathbf{e}_{tk}\mathbf{e}_{tk}'\right)\right\} p(\Sigma_k) \tag{A.28}$$

where $\mathcal{T}_k = \{t : s_t = k\}$, $\mathbf{e}_t = \mathbf{y}_t - Z_t'\boldsymbol{\beta}(s_t)$ and $\mathbf{e}_{tk} = \mathbf{y}_t - Z_t'\boldsymbol{\beta}_k$. Exploiting the conditional independence structure encoded in $k$-th state graph $G_k$ it follows

$$\begin{aligned}
&p\left(\Sigma_k|\mathbf{y}_{1:T}, \boldsymbol{\theta}, \mathbf{s}_{1:T}, \boldsymbol{\beta}_k\right) \propto \\
&\propto \prod_{j=1}^{n_P} \left|\Sigma_{P_j,k}\right|^{-T_k/2} \exp\left\{-\frac{1}{2}\Sigma_{P_j,k}^{-1}A_{P_j,k}^*\right\} \times \prod_{j=1}^{n_P} \left|\Sigma_{P_j,k}\right|^{-\left(a_k+2T_{P_j,k}\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{P_j,k}^{-1}A_{P_j,k})\right\} \cdot \\
&\quad \cdot \prod_{j=1}^{n_B} \left(\left|\Sigma_{B_j,k}\right|^{-\left(a_k+2T_{B_j,k}\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{B_j,k}^{-1}A_{S_j,k})\right\}\right)^{-1} \\
&\propto \prod_{j=1}^{n_P} \left|\Sigma_{P_j,k}\right|^{-\left(a_k+2T_{P_j,k}+T_k\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{P_j,k}^{-1}(A_{P_j,k}+A_{P_j,k}^*)\right\} \cdot \\
&\quad \cdot \prod_{j=1}^{n_B} \left(\left|\Sigma_{B_j,k}\right|^{-\left(a_k+2T_{B_j,k}\right)/2} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma_{B_j,k}^{-1}A_{B_j,k})\right\}\right)^{-1} \\
&\propto \mathcal{HIW}_{G_k}\left(a_k + T_k, A_k + A_k^*\right),
\end{aligned} \tag{A.29}$$

where $A^*_{P_j,k}$ is the block of $A^*_k$ corresponding to $\Sigma_{P_j,k}$, $T_{B_j,k} = Card(B_j, k)$ and $\Sigma_{B_j,k}$ defined over the set of separators as Eq. (9).

## A.2 Proof of Proposition 1

The proof of Proposition 1 is in two parts. First we show that our weighted eigenvector centrality measure is the limit of the rescaled Bonacich's $c(\beta)$centrality (see Bonacich 2007) of the graph $\tilde{G}_k$ of regime $k \in K$. In the second part we show that the Bonacich's $c(\beta)$ centrality can be written as a state-specific weighed sum over all walks between all pairs of nodes of the graph $\tilde{G}_k$.

Let $\tilde{\sigma}_{ij,k}$ be the $i$-th row $j$-th column entry of $\tilde{\Sigma}_k$, the relative weighted centrality scores of $\tilde{G}_k$

$$\lambda_k \gamma_{i,k} = \sum_{j=1}^{n} \tilde{\sigma}_{ij,k} \gamma_{j,k}, \quad i = 1, \ldots, n$$

can be re-written in a more compact format as the eigenvector equation $\tilde{\Sigma}_k \boldsymbol{\gamma}_k = \lambda_k \boldsymbol{\gamma}_k$ with $n$ solutions given by the eigenvalues $\lambda_{i,k}$, with $\lambda_{1,k} \leq \ldots \leq \lambda_{n,k}$, and the associated eigenvectors $\boldsymbol{\gamma}_k = (\gamma_{1,k}, \ldots, \gamma_{n,k})'$. Since $\tilde{\Sigma}_k$ is symmetric then the decomposition $\tilde{\Sigma}^l_k = \boldsymbol{\gamma}_k \Lambda^l_k \boldsymbol{\gamma}'_k$ holds true with $\Gamma_k = (\boldsymbol{\gamma}_{1,k}, \ldots, \boldsymbol{\gamma}_{n,k})$, $\Lambda_k = \text{diag}\{\lambda_{1,k}, \ldots, \lambda_{n,k}\}$, and $\gamma_{i,k}$ $i = 1, \ldots, n$ orthonormal eigenvectors.

In general, the Bonacich's $c(\beta)$ centrality can be defined as

$$c(\beta) = \lim_{L \to \infty} \sum_{l=1}^{L} \beta^{l-1} \tilde{\Sigma}^l_k \boldsymbol{\iota}$$

with $\boldsymbol{\iota}$ the $n$-dimensional unit vector. By applying the previous decomposition this measure can be written as

$$\begin{aligned} c(\beta) &= \lim_{L \to \infty} \frac{1}{\beta} \sum_{l=1}^{L} \beta^l (\boldsymbol{\gamma}_k \Lambda^l_k \boldsymbol{\gamma}'_k) \boldsymbol{\iota} = \lim_{L \to \infty} \frac{1}{\beta} \sum_{l=1}^{L} \beta^l \left( \sum_{i=1}^{n} \gamma_{i,k} \lambda^l_{i,k} \gamma'_{i,k} \right) \boldsymbol{\iota} = \\ &= \frac{1}{\beta} \sum_{i=1}^{n} \left( \lim_{L \to \infty} \sum_{l=1}^{L} \beta^l \lambda^l_{i,k} \right) \gamma_{i,k} \gamma'_{i,k} \boldsymbol{\iota} \end{aligned} \tag{A.30}$$

Under the assumptions the eigenvectors are orthonormal and $|\beta| < 1/\lambda_{n,k}$ it follows that (see, e.g. Bonacich 2007);

$$c(\beta) = \frac{1}{\beta}\sum_{i=1}^{n}\varphi_i(\beta,k)\gamma_{i,k},$$

where

$$\varphi_i(\beta,k) = \frac{\beta\lambda_{i,k}}{1-\beta\lambda_{i,k}}$$

Now considering the limit as $\beta$ tends to $1/\lambda_{n,k}^{-}$

$$\lim_{\beta\to 1/\lambda_{n,k}^{-}}\left(\sum_{i=1}^{n}\varphi_i(\beta,k)\right)^{-1}c(\beta) = \sum_{i=1}^{n}\left(\lim_{\beta\to 1/\lambda_{n,k}^{-}}\frac{\varphi_i(\beta,k)}{\sum_{j=1}^{n}\varphi_j(\beta,k)}\gamma_{i,k}\right) = \gamma_{n,k} \qquad \text{(A.31)}$$

since $\varphi_n(\beta,k)$ diverges with $\beta$ and $\varphi_i(\beta,k)$ converges to a finite quantity for all $i < n$. To conclude the first part of the proof we just need to find the relationship between our centrality measure $\overline{q}(\tilde{G}_k)$ and the limit given above. This follows immediately from the definition of $\overline{q}(\tilde{G}_k)$, that is

$$\overline{q}(\tilde{G}_k) = \frac{1}{n}\boldsymbol{\iota}'\gamma_{n,k} = \lim_{\beta\to\lambda_{n,k}^{-}}\frac{1}{\kappa(\beta,k)}\boldsymbol{\iota}'c(\beta) \qquad \text{(A.32)}$$

where $\kappa(\beta,k) = \frac{1}{n}\sum_{j=1}^{n}\varphi_j(\beta,k)$.

For the second part of the proof, we can use Definition 1 and 2 in the main text and rewrite part the right-hand side of Eq. A.32 as:

$$\boldsymbol{\iota}'c(\beta) = \lim_{L\to\infty}\frac{1}{\beta}\sum_{l=1}^{L}\beta^l\boldsymbol{\iota}'\tilde{\Sigma}_k^l\boldsymbol{\iota} \qquad \text{(A.33)}$$

which can be further re-written as a function of the sum of the weights over all possible

walks between all pairs of nodes;

$$\boldsymbol{\iota}'\tilde{\Sigma}_k^l\boldsymbol{\iota} = \sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\sigma}_{ij,k}^{(l)} = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{i_1\in V_k}\tilde{\sigma}_{ii_1,k}^{(l-1)}\tilde{\sigma}_{i_1j,k} \tag{A.34}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{i_1\in V_k}\sum_{i_2\in V_k}\tilde{\sigma}_{ii_2,k}^{(l-2)}\tilde{\sigma}_{i_2i_1,k}\tilde{\sigma}_{i_1j,k} = \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{i_1\in V_k}\cdots\sum_{i_{l-1}\in V_k}\prod_{r=1}^{l}\tilde{\sigma}_{i_{r-1}i_r,k}\right) \tag{A.35}$$

Let $a_{i_{r-1}i_r,k}$ be the $(i_{r-1},i_r)$-th element of the adjacency matrix $A_k$ associated with $G_k$ and $P_{ij,k}^{(l)} = \{(D_k(p),V_k(p));\, V_k(p)=\{i_0,\ldots,i_l\}\subset V_k,\, D_k(p)=\{e_1,\ldots,e_l\}\subset E, i_0 = i, i_l = j\}$ be the set of all walks of length $l$ between $i$ and $j$. Since $\prod_{r=1}^{l}\tilde{\sigma}_{i_{r-1}i_r,k} = \prod_{r=1}^{l}a_{i_{r-1}i_r,k}\sigma_{i_{r-1}i_r,k}$ and

$$\prod_{r=1}^{l}a_{i_{r-1}i_r,k} = \begin{cases} 1 & \text{if } (i_0,e_1,\ldots,e_l,i_l)\in P_{ij,k}^{(l)} \\ 0 & \text{otherwise} \end{cases}$$

it follows that

$$\boldsymbol{\iota}'\tilde{\Sigma}_k^l\boldsymbol{\iota} = \sum_{i,j\in V}\sum_{p\in P_{ij,k}^{(l)}}\prod_{r=1}^{l}\sigma_{i_{r-1}i_r,k} \tag{A.36}$$

where $\prod_{r=1}^{l}\sigma_{i_{r-1}i_r,k}$ is the path weight. By plugging (A.36) in (A.33) and substituting the result in (A.32) we obtain Eq.(21) in the text.

## A.3 Proof of Corollary 1

From Proposition 1, our centrality measure has the following representation

$$\begin{aligned}\overline{q}(\tilde{G}_k) &= \lim_{\beta\to 1/\lambda_{n,k}^-}\frac{1}{\kappa(\beta,k)}\sum_{l=1}^{\infty}\beta^l\boldsymbol{\iota}'\tilde{\Sigma}_k^l\boldsymbol{\iota} \\ &= \lim_{\beta\to 1/\lambda_{n,k}^-}\frac{1}{\kappa(\beta,k)}\sum_{i,j\in V_k}\sum_{l=1}^{\infty}\beta^l\sum_{p\in P_{ij,k}^{(l)}}\prod_{r=1}^{l}\sigma_{i_{r-1}i_r,k}\end{aligned} \tag{A.37}$$

Then, since it is possible to obtain a path from a walk by removing its sub-cycles we can define for each pair of nodes $i$ and $j$ the set of paths $P_{ij,k}^{*}$ between $i$ and $j$ with

minimum length $l^*_{ij}$ (shortest paths) and the set $P^{(l)}_{ij,k}$ of all remaining walks between the two nodes with length $l > l^*_{ij}$. Since $P^{(l)}_{ij,k} = \emptyset$ for all $l < l^*$, it follows that

$$\sum_{l=1}^{\infty} \beta^l \sum_{p \in P^{(l)}_{ij,k}} \prod_{r=1}^{l} \sigma_{i_{r-1} i_r, k} = \left( \sum_{p \in P^{*}_{ij,k}} \beta^{l^*_{ij}-1} \prod_{r=1}^{l} \sigma_{i_{r-1} i_r, k} + \sum_{l > l^*_{ij}}^{\infty} \beta^{l-1} \sum_{p \in P^{(l)}_{ij,k}} \prod_{r=1}^{l} \sigma_{i_{r-1} i_r, k} \right)$$

which gives the desired decomposition.

## A.4 Proof of Corollary 2

The weighted eigenvector centrality measure $\overline{q}(\tilde{G}_k)$ can be written as

$$\overline{q}(\tilde{G}_k) = \lim_{\beta \to 1/\lambda^-_{n,k}} \frac{1}{\kappa(\beta, k)} \frac{1}{\beta} \sum_{i=1}^{n} \sum_{l=1}^{\infty} (\beta \lambda_{i,k})^l \boldsymbol{\iota}' \boldsymbol{\gamma}_{i,k} \boldsymbol{\gamma}'_{i,k} \boldsymbol{\iota} \tag{A.38}$$

Let us denote the element-wise positive and negative parts of $\boldsymbol{\gamma}_{i,k}$, with $\boldsymbol{\gamma}^+_{i,k}$ and $\boldsymbol{\gamma}^-_{i,k}$ respectively. Then, $\boldsymbol{\iota}' \boldsymbol{\gamma}_{i,k} = \boldsymbol{\iota}' (\boldsymbol{\gamma}^+_{i,k} - \boldsymbol{\gamma}^-_{i,k})$ and

$$\sum_{i=1}^{n} \frac{\beta \lambda_{i,k}}{1 - \beta \lambda_{i,k}} \boldsymbol{\iota}' \boldsymbol{\gamma}_{i,k} \boldsymbol{\gamma}'_{i,k} \boldsymbol{\iota} = \sum_{i=1}^{n} \frac{\beta \lambda_{i,k}}{1 - \beta \lambda_{i,k}} (\boldsymbol{\iota}' \gamma^+_{i,k})^2 + \sum_{i=1}^{n} \frac{\beta \lambda_{i,k}}{1 - \beta \lambda_{i,k}} (\boldsymbol{\iota}' \boldsymbol{\gamma}^-_{i,k})^2$$
$$- 2 \sum_{i=1}^{n} \frac{\beta \lambda_{i,k}}{1 - \beta \lambda_{i,k}} \boldsymbol{\iota}' \gamma^+_{i,k} \boldsymbol{\iota}' \gamma^-_{i,k}$$

which yields the decomposition for $\overline{q}(\tilde{G}_k)$ given in Eq. (23).

# B Simulated Evidence

In this section, we investigate the ability of our inferential scheme to detect the existence of linkages in a network of moderate size, that represents our null model, i.e., from which we perform simulations. To keep simulations and estimations simple, we consider a single-factor model. We also compare our Markov switching Gaussian graphical model to a similar specification without the graph structure in the covariance matrix of residuals. We simulate a sample of $T = 1000$ observations $\mathbf{y}_t$, for $n = 20$ assets and considering a single factor $Z_t \sim i.i.d.\ \mathcal{N}(0, 1)$. For simplicity, we

assume that the betas on a single factor are constant across assets and are different across states, $\beta_{i,1} = 0.6$ and $\beta_{i,2} = 1.2$, for $i = 1, \ldots, n$, as well as the existence of two persistent states with $\pi_{11} = \pi_{22} = 0.95$. The residual covariance structure also changes across regimes and is consistent with an underlying graph-based network $G_k$. Network connectedness is set to be more concentrated in state $s_t = 2$, which therefore represents a regime of high systemic risk. To avoid any particular effect of prior elicitation we choose fairly vague priors for both states. Top Panels of Figure A.1 show the adjacency matrix that defines the true network against the posterior median estimates of $G_2$;

[Insert Figure A.1 about here]

The figure makes it clear that the model has a fairly good performance in identifying network connectivity; visually, the estimated network almost entirely overlaps with the original. We further compare our Markov switching graphical model with a benchmark without network structure in the residuals, i.e. $\mathcal{M}_2$, by computing the estimation risk on $\Sigma_k$ for different sample sizes using Stein's loss function

$$L\left(\hat{\Sigma}_2, \Sigma_2\right) = tr\left(\hat{\Sigma}_2 \Sigma_2^{-1}\right) - \log|\hat{\Sigma}_2 \Sigma_2^{-1}| - n \tag{A.39}$$

where $\hat{\Sigma}_2$ and $\Sigma_2$ are the posterior median estimates and true covariance matrix of the residuals assuming $s_t = 2$, respectively. We conduct the experiment for different sample sizes, $T = 50, 100$, and 200, with $n = 20$ assets and considering a single factor as above, under identical specifications for the betas and state persistence. Bottom Panel of Figure A.1 shows box plots of the risk associated by different estimators across different sample sizes. The figure shows that our model, i.e. $\mathcal{M}_1$, offers large gain over a standard Markov-switching specification. This gain is particularly significant when the ratio between the number of assets and the sample size $(n/T)$ is relatively large. This is consistent with previous evidence on the efficiency of sparse covariance estimators (see e.g. Carvalho et al. 2007 and Wang and West 2009, among others).

# C Testing the Number of Regimes

Our Markov regime switching Gaussian graphical model implies that network connectivity is state-dependent with a finite number of regimes. Economic theory assumes that contagion and systemic risk represent more of a shift concept than steady, recurring regimes justifying our initial choice of $K = 2$. As a result, one may argue that two states may not enough to capture the dynamic features of an extensive and complex network such as the one that is likely defined by US equities. To deal with this conjecture, we test the null hypothesis $\mathcal{H}_0 : K = 2$ against the alternative $\mathcal{H}_1' : K = 1$ and $\mathcal{H}_1'' : K = 3$ on the basis of Bayes factors comparing the model with $i$ regimes $\mathcal{M}_i$ against a model with $j$ regimes $\mathcal{M}_j$. Bayes factors are based on marginal likelihoods (see Kass and Raftery 1995) so that comparing, a two-state vs. a $j$-state model can be accomplished by computing:

$$\mathcal{B}_{ij} = \frac{p(\mathbf{y}_{1:T}|\mathcal{M}_i)p\left(\mathcal{M}_i\right)}{p(\mathbf{y}_{1:T}|\mathcal{M}_j)p\left(\mathcal{M}_j\right)}, \tag{A.40}$$

As customary, values of $\mathcal{B}_{i,j}$ in excess of 10, or better values of $\log_{10}\mathcal{B}_{i,j}$ in excess of 1 are to be considered strong evidence in favor of $\mathcal{M}_i$ over $\mathcal{M}_j$. In our application, the marginal likelihoods are computed by integrating out both parameter and state uncertainty from the posterior distribution of latent states and parameters obtained from the Metropolis-within-Gibbs sampler detailed in section 3. The marginal likelihood however, is not available in closed form and can be approximated numerically as in Chib (1995). We assume $p\left(\mathcal{M}_i\right) = p\left(\mathcal{M}_j\right)$. Table C.1 shows the results across models.

[Insert Table C.1 about here]

Panel A shows that the marginal likelihoods are systematically higher for models with two regimes, meaning that the empirical evidence provided by Bayes factors in log-10 scale is strongly in favor of a model with two regimes vs. three states and two vs. time-invariant models. In fact the minimal value for $\log_{10}\mathcal{B}_{2,3}$ across different pricing models is 6.22 while the minimal value for $\log_{10}\mathcal{B}_{2,1}$ is 6.16. This evidence should not be subject to dispute.

## **Table 1.** Network Centrality and Market Values

This table reports the results from a cross-sectional regression where the dependent variable is the centrality measure computed for each firm and industry and the explanatory variable is the firm-specific market value. We control for industry heterogeneity by including and industry fixed-effect for both regimes of systemic risk. Kendall (1938) rank correlation coefficients are computed by first ranking firms according to their centrality within the network, then ranking firms according to their average market value for each regime. The rank correlation coefficient $\tau$ measures the correspondence of the ranking. Standard errors are corrected for heteroskedasticity and autocorrelation in the residuals using Newey-West HAC correction. Rank correlations are highlighted in grey when the null hypothesis is rejected at least at a 5% significance level.

| Weighted Eigenvector Centrality | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CAPM | | | | Fama-French | | | | I-CAPM | | | |
| | Coeff | t-stat | $R^2$ | $\tau$ | Coeff | t-stat | $R^2$ | $\tau$ | Coeff | t-stat | $R^2$ | $\tau$ |
| High | 0.021 | 1.201 | 0.012 | 0.054 | 0.011 | 1.291 | 0.021 | 0.061 | 0.021 | 1.254 | 0.013 | 0.045 |
| Low | 0.035 | 1.581 | 0.029 | 0.084 | 0.052 | 1.271 | 0.013 | 0.085 | 0.039 | 1.432 | 0.024 | 0.055 |

## **Table 2.** Network Centrality and Realized Financial Losses

This table reports the results from a cross-sectional regression model where the dependent variable is the ranking of firms (within industries) on the basis of their average maximum percentage financial loss suffered across the two separate regimes and the explanatory variable is the firm-specific market value. We control for industry heterogeneity by including and industry fixed-effect for both regimes of systemic risk. Kendall (1938) rank correlation coefficients are computed by first ranking firms according to their centrality within the network, then ranking firms according to their average market value for each regime. The rank correlation coefficient $\tau$ measures the correspondence of the ranking. Standard errors are corrected for heteroskedasticity and autocorrelation in the residuals using the Newey-West HAC correction. Rank correlations are highlighted in grey when the null hypothesis is rejected at least at a 5% significance level.

| Panel A: Weighted Eigenvector Centrality | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CAPM | | | | Fama-French | | | | I-CAPM | | | |
| | Coeff | t-stat | $R^2$ | $\tau$ | Coeff | t-stat | $R^2$ | $\tau$ | Coeff | t-stat | $R^2$ | $\tau$ |
| High | 0.751 | 2.261 | 0.121 | 0.211 | 0.871 | 2.314 | 0.101 | 0.205 | 0.340 | 2.181 | 0.112 | 0.198 |
| Low | 0.412 | 1.901 | 0.145 | 0.181 | 0.301 | 1.913 | 0.062 | 0.171 | 0.456 | 1.859 | 0.061 | 0.169 |

## **Table C.1.** Testing the Number of Regimes

This table reports the results of a formal test for the number of regimes for each factor pricing model specification. We report the (log) marginal likelihoods and the corresponding Bayes factor in log-scale comparing the null hypothesis $\mathcal{H}_0 : K = 2$ against the alternative $\mathcal{H}_1' : K = 1$ and $\mathcal{H}_1'' : K = 3$. Bayes factors are based on marginal likelihoods computed by integrating out both parameter and state uncertainty from the posterior distribution of latent states and parameters obtained from the Metropolis-within-Gibbs sampler (see Kass and Raftery 1995).

| Panel A: Log Marginal Likelihoods | | | |
|---|---|---|---|
| | CAPM | Fama-French | I-CAPM |
| $K = 1$ | $-1.84e+06$ | $-1.86e+06$ | $-1.87e+06$ |
| $K = 2$ | $-1.75e+05$ | $-1.63e+05$ | $-1.56e+05$ |
| $K = 3$ | $-1.60e+06$ | $-1.62e+06$ | $-1.63e+06$ |

| Panel B: $Log_{10}$ Bayes Factors | | | |
|---|---|---|---|
| | CAPM | Fama-French | I-CAPM |
| $\mathcal{B}_{k=2,k=3}$ | 6.2223 | 6.2302 | 6.2346 |
| $\mathcal{B}_{k=2,k=1}$ | 6.1547 | 6.1632 | 6.1680 |

**Figure 1.** Systemic Risk Filtered Probabilities

This figure shows the filtered probability of being in a state of high systemic risk computed from the Markov Switching Gaussian graphical model. The left panel shows the probabilities from the Fama-French three-factor specification, and the right panel shows the results from an I-CAPM. The grey area represents the filtered probability of high-systemic risk, the red solid line shows the NBER recession indicator for the period following the peak of the recession to the trough.

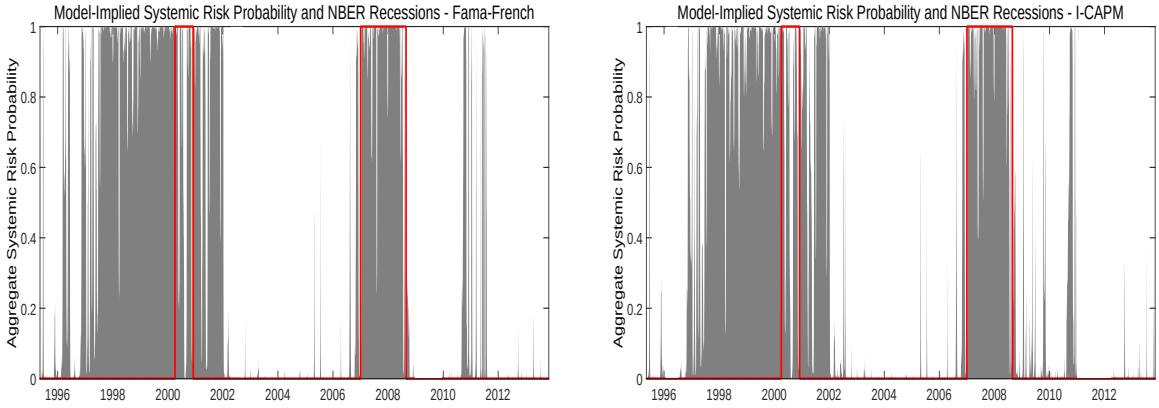



**Figure 2.** Transition Probabilities

This figure shows the transition probabilities of the latent states from the Markov switching Gaussian graphical model. The first (last) three columns represent the probability of staying in a state of low (high) systemic risk. The red line shows the posterior median, the blue box reports the $25th$ and $75th$ quantiles.

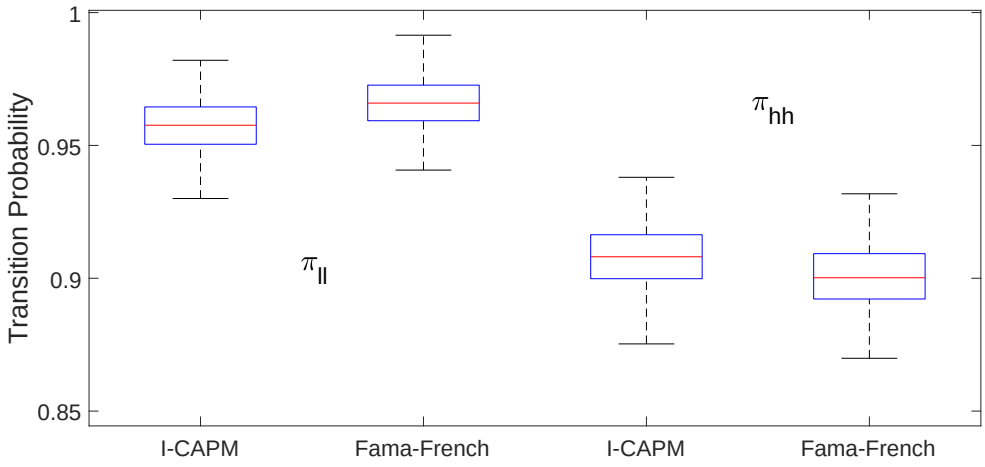


**Figure 3.** Changes in Betas Across Regimes, Three-Factor Model

This figure shows the changes in the intercepts (to be interpreted as a Jensen's alpha because of the tradability of the factors) and betas between two regimes of low and high systemic risk, respectively, and across stocks for the Fama-French three-factor model. The top left panel shows the changes in the interecept, i.e. Jensen's alpha. The top right panel reports the changes in the betas on the market portfolio. The bottom left and right panels report the changes in the betas measuring exposures to size and value mimicking portfolios.

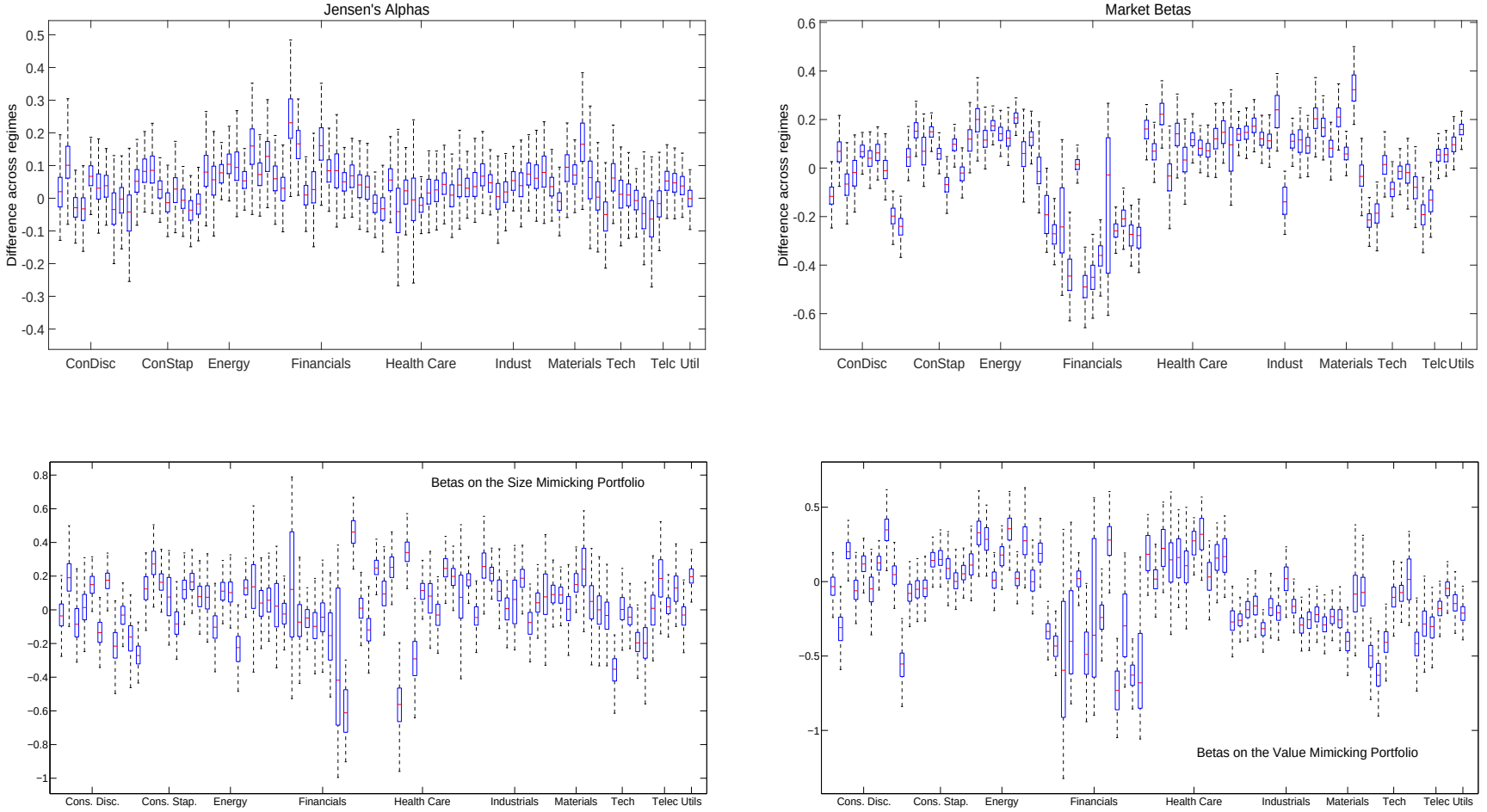


**Figure 4.** Changes in Betas Across Regimes, I-CAPM

This figure shows the changes in the intercepts (to be interpreted as a Jensen's alpha because of the tradability of the factors) and betas between two regimes of low and high systemic risk, respectively, and across stocks for an I-CAPM. The top left panel shows the changes in the interecept, i.e. Jensen's alpha. The top right panel reports the changes in the betas on the market portfolio. The bottom left and right panels report the changes on the betas measuring exposure to default risk and aggregate dividend yield.

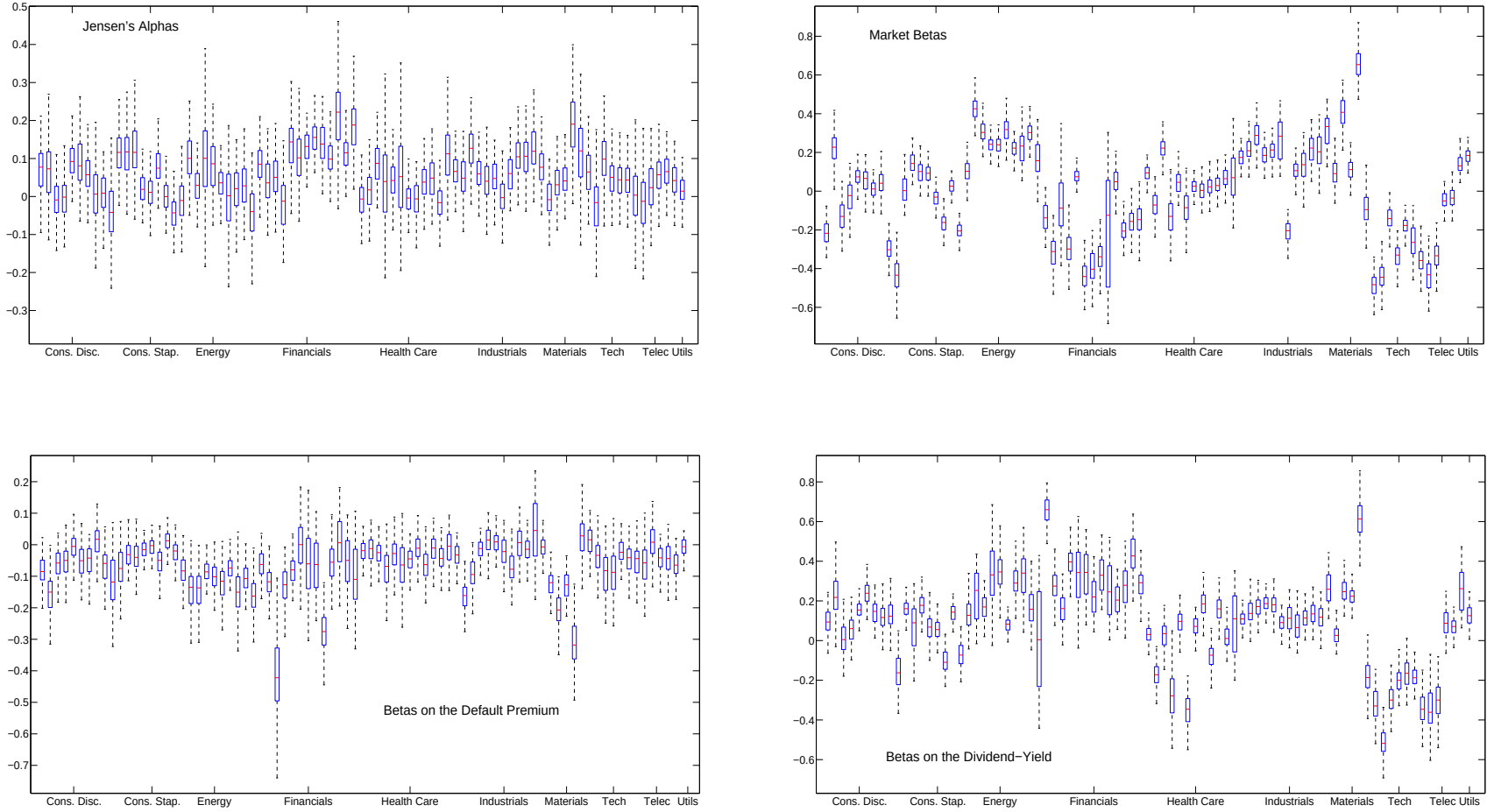
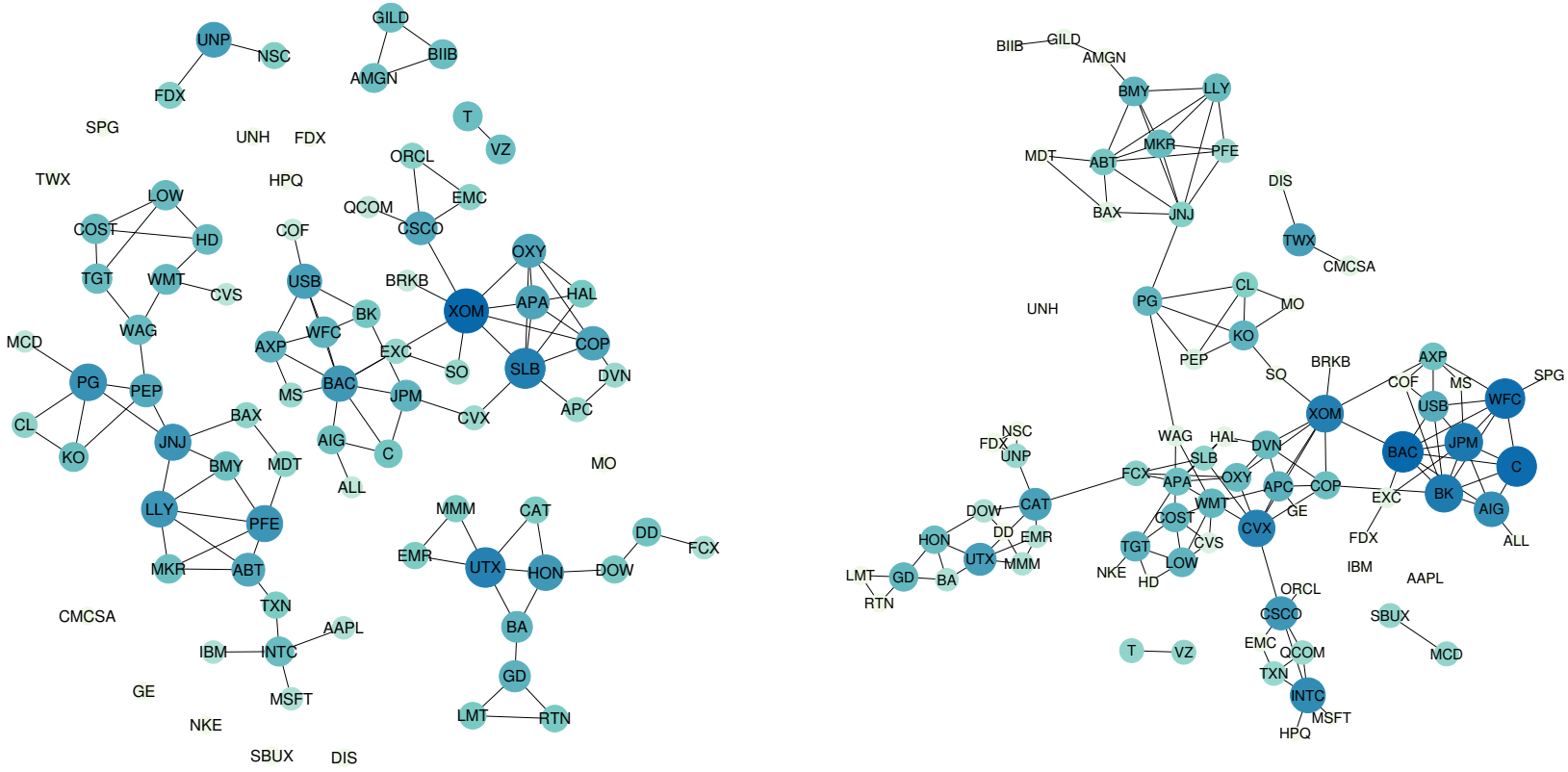


**Figure 5.** Network Connectivity: Three-Factor Model

This figure reports the posterior median estimates of the network across regimes obtained from the Fama-French three-factor model residuals. The left panel shows the structure of the network when cross-firm connectedness is low; the right panel shows the structure of the network when cross-firm connectedness is high. The size and the color of the nodes are proportional to their relevance in the network measured by weighted eigenvector centrality. The darker (bigger) the color (size) of the node, the higher its marginal contribution to aggregate systemic risk.

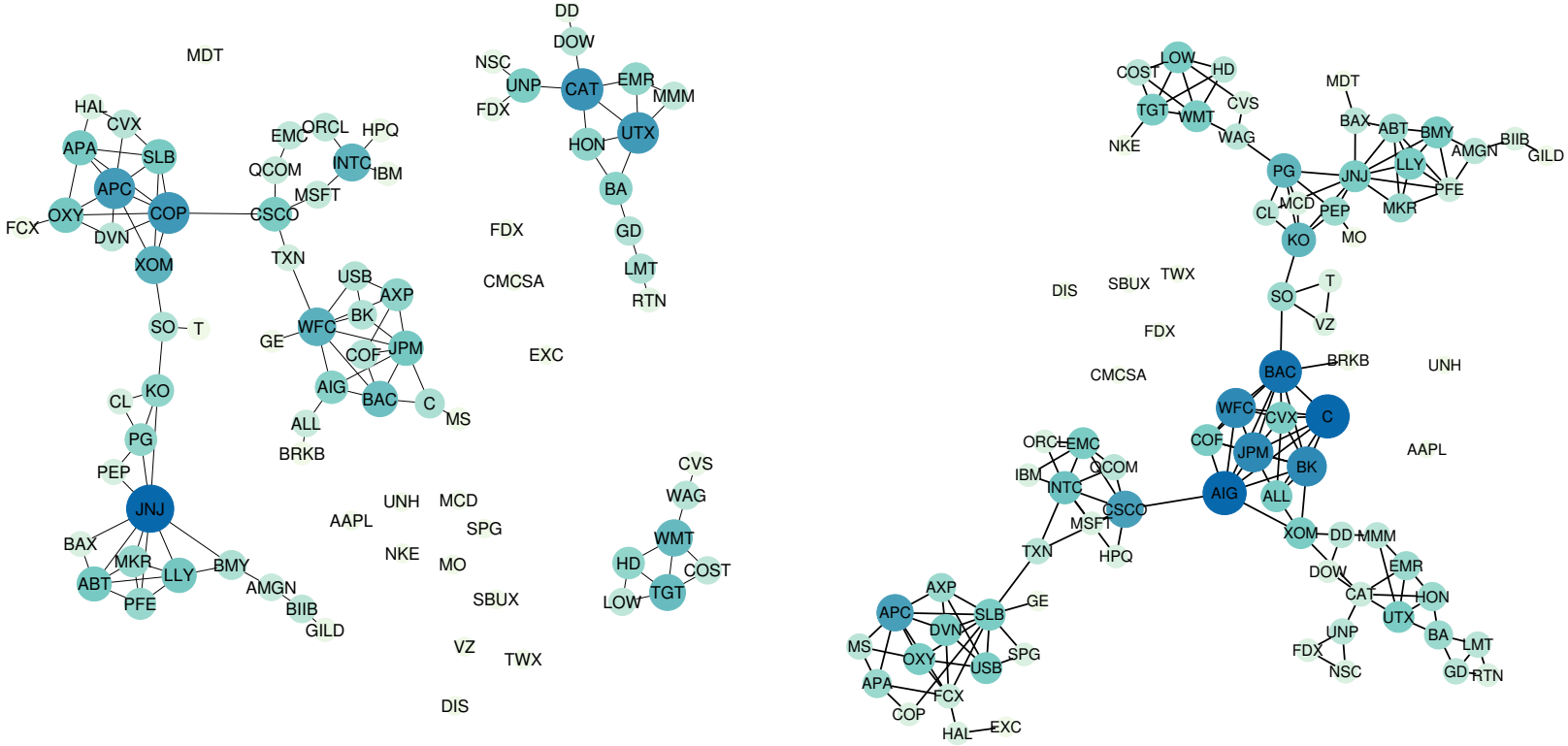



(a) Low Systemic Risk

(b) High Systemic Risk

**Figure 6.** Network Connectivity: I-CAPM

This figure reports the posterior median estimates of the network across regimes obtained from the I-CAPM model residuals. The left panel shows the structure of the network when cross-firm connectedness is low; the right panel shows the structure of the network when cross-firm connectedness is high. The size and the color of the nodes are proportional to their relevance in the network measured by weighted eigenvector centrality. The darker (bigger) the color (size) of the node, the higher its marginal contribution to aggregate systemic risk.




(a) Low Systemic Risk

(b) High Systemic Risk

**Figure 7.** Firm-Level Network Centrality

This figure plots the posterior median weighted eigenvector centrality sorted for the top 20 stocks in both low (red line) and high (blue line) systemic risk. The network structure is estimated from the residuals of a Fama-French three-factor model (left panel) and an I-CAPM (right panel). Graphs are sampled from the Metropolis-within-Gibbs sampler outlined in Section 3.

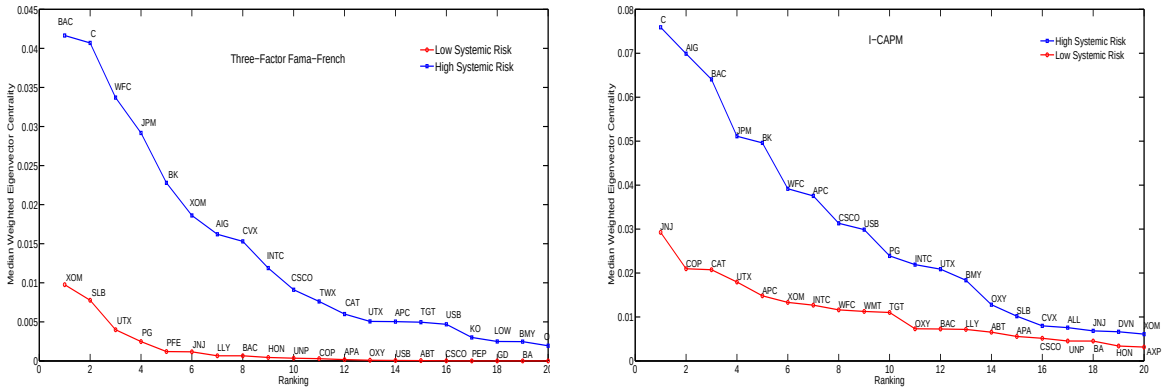



**Figure 8.** Industry-Level Network Centrality

This figure plots the posterior median weighted eigenvector centrality clustered at the industry level. The industry-level centrality measures are obtained by taking the median of firm-specific measures averaged out within industries. Posterior median estimates of the network structure is computed from the residuals of a Fama-French three-factor model (green line) and an I-CAPM specification (blue line). Graphs are sampled from the Metropolis-within-Gibbs sampler outlined in Section 3.

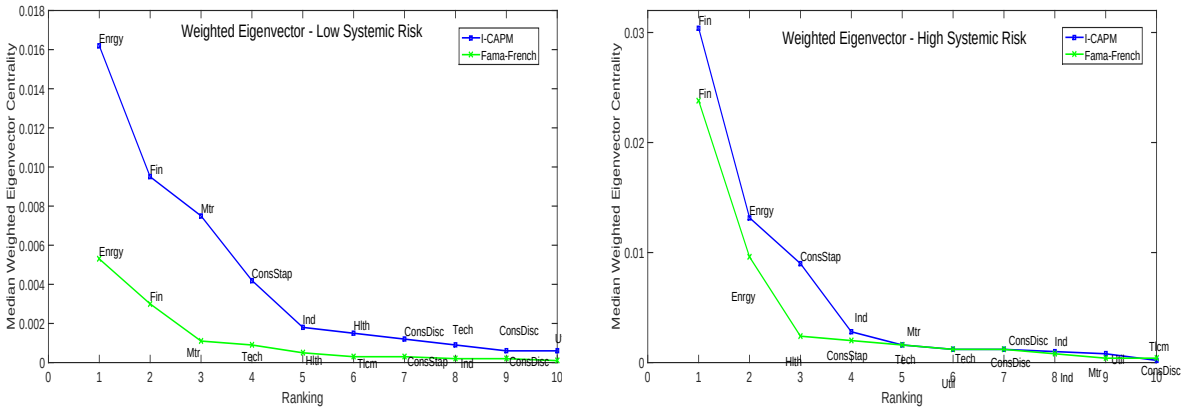


**Figure A.1.** Simulation Example

This figure plots the estimation results on a simulated dataset. Top panels compare the estimated network (right) to the true one (left). The length of the simulated time series is $T = 1000$ and the number of unites (assets) is $n = 20$. Bottom panel shows the results of applying a Stein Loss to compare our model to a standard Markov regime-switching SUR model without network structure in the residuals and for different sample sizes equal to $T = 50, 100, 200$. We consider $n = 20$ and a single factor as above.