

## Ripples on financial networks

Sudarshan Kumar  
Avijit Bansal  
Anindya S. Chakrabarti<sup>a</sup>

**W.P. No. 2019-10-01**

October 2019

---

<sup>a</sup>(corresponding author) anindyac@iima.ac.in. Economics area, Indian Institute of Management, Vastrapur, Ahmedabad, Gujarat-380015, India.

The main objective of the Working Paper series of IIMA is to help faculty members, research staff, and doctoral students to speedily share their research findings with professional colleagues and to test out their research findings at the pre-publication stage.

**INDIAN INSTITUTE OF MANAGEMENT  
AHMEDABAD – 380015  
INDIA**

RIPPLES ON FINANCIAL NETWORKS<sup>1</sup>

Sudarshan Kumar<sup>2</sup>  
Avijit Bansal<sup>3</sup>  
Anindya S. Chakrabarti<sup>4</sup>

**Abstract**

In the financial markets, asset returns exhibit collective dynamics masking individual impacts on the rest of the market. Hence, it is still an open problem to identify how shocks originating from one particular asset would create spillover effects across other assets. The problem is more acute when there is a large number of simultaneously traded assets, making the identification of which asset affects which other assets even more difficult. In this paper, we construct a network of the conditional volatility series estimated from asset returns and propose a many-dimensional VAR model with unique identification criteria based on the network topology. Because of the interlinkages across stocks, volatility shock to a particular asset propagates through the network creating a ripple effect. Our method allows us to find the exact path the ripple effect follows on the whole network of assets.

**Keywords:** Volatility spillover, VAR identification, Spectral analysis, Core-periphery

**JEL Codes:** C55, C58, G17, C53.

---

<sup>1</sup>This research was partially supported by institute grant, IIM Ahmedabad. All remaining errors are ours.

<sup>2</sup>Email: sudarshank@iima.ac.in. Finance & accounting area, Indian Institute of Management, Vastrapur, Ahmedabad, Gujarat-380015, India.

<sup>3</sup>Email: fpm16avijitb@iima.ac.in. Finance & accounting area, Indian Institute of Management, Vastrapur, Ahmedabad, Gujarat-380015, India.

<sup>4</sup>(corresponding author) anindyac@iima.ac.in. Economics area, Indian Institute of Management, Vastrapur, Ahmedabad, Gujarat-380015, India.

# 1 Introduction

Financial networks representing concurrent evolution of asset returns and volatilities across time, show non-trivial aggregate dynamics that arise out of individual dynamic of all underlying assets. Assessing contributions of individual assets to the aggregate behavior is difficult because assets returns do not evolve over time in silos. They are affected by and in turn affect other asset returns. Thus, quantification of the effects of a volatility shock emanating from a given epicenter percolating through the network would depend on the topology of the network denoting linkages across asset returns as well as the shock structure. In this paper, we propose an algorithm to characterize such *ripples* on large-scale financial networks.

There has been considerable development in the theoretical and empirical literature in the aftermath of the recent US financial crisis which brought forward the importance of connections across economic and financial entities (Acemoglu et al. (2015), Acemoglu et al. (2012), Barigozzi and Hallin (2017b), Dessaint et al. (2018), Diebold and Yilmaz (2015a), Pozzi et al. (2013), Tumminello et al. (2010), Barigozzi and Hallin (2017a), Geraci and Gnabo (2018), Tumminello et al. (2007)). The idea that spillover effects can be of sizable magnitude came to the forefront almost immediately as has been demonstrated by the remark by the then chief executive of Ford Motors, A. Mullaly in the congressional hearing (Mullaly (2008)) during the crisis: *“If any one of the domestic companies should fail, we believe there is a strong chance that the entire industry would face severe disruption.”* One stream of the literature took the path of understanding the shock propagation mechanism through input-output architecture (see Acemoglu et al. (2016) for a review of the literature). In the financial markets however, the shock propagation channel can take many forms. Plosser (2009) noted that: *“due to the complexity and interconnectivity of today’s financial markets, the failure of a major counterparty has the potential to severely disrupt many other financial institutions, their customers, and other markets.”* We focus on the second stream of literature which attempts to address such phenomena from a statistical point of view where one has to construct and/or estimate the network from underlying process and data. Essentially, we do not provide a structural reason for shock propagation. This is in line with the reduced-form approach taken by Demirer et al. (2017).

Modeling spillover of shocks across a large number of stocks presents two challenges. First, since all stocks are traded at a very high frequency, the daily data reflect the collective dynamics and each individual return series comprise own dynamics as well as impacts of other assets’ dynamics. Thus identification of which asset affects which other assets cannot be deduced by considering stocks separately. We accommodate the information of joint evolution by constructing network of stocks and for modeling the time domain information, we utilize a vector autoregression (in short, VAR) framework with an identification restriction obtained from the underlying network topology. This leads to the second problem about the robustness of the identification criteria. In particular, given the rapid evolution in the financial markets, finding a criteria that produces time-consistent result is difficult. In the following, we show that the criteria that we derive from topology of the network based on observed return correlation matrix is quite stable across time.

In terms of methodology, we first construct the conditional volatility series from stock return data using a generalized autoregressive conditional heteroschedastic (in short, GARCH) framework. We allow the lags to vary between a range and by using information criteria, we select the best lag combinations for each of the stock return series. Since the number of stocks are large and not all pairs of stocks would be strongly connected, we carry out a sample selection procedure by fitting a sparse VAR model (Lasso estimator; Tibshirani (1996)) to the full set of stocks volatility series to identify a subset of strongly connected stocks

using an exogenously tunable threshold. The surviving set of stocks constitute the *maximally connected component*.

Once the set of connected stocks is identified we construct a network based on observed return correlation matrix. Since the correlation values can be negative, it does not satisfy the criteria for a metric. Hence, the correlation values cannot directly be used for creating a network. Following Mantegna (1999), we take a nonlinear transformation (see Sec. 2.2.2) to convert the correlation matrix into a distance matrix. Then we extract the minimum spanning tree (MST herewith) from the same, which is ultra-metric in nature (Mantegna (1999)). The MST provides the minimal structure on which we demonstrate the shock propagation mechanism in the form of ripple effects.

In order to analyze shock propagation, we estimate a VAR model on the stocks belonging to the maximally connected component, with suitable lags chosen by information criteria. The model is identified by imposing an ordering on the matrix representing contemporaneous movement of the volatility variables, by the eigenvector centrality of the observed return correlation matrix. This identification choice reflects two mechanisms. First, the dominant eigenvalue of the observed return correlation matrix represents the so-called *market factor* (Plerou et al. (2002)). The name arises from the fact that the corresponding eigenvector contains the extent of participations of the stocks in the aggregate movement. Thus a high value of a stock in the dominant eigenvector represents higher participation of the stock, which provides information about relative exogeneity of the assets in the market. Since volatility spillovers will occur from a more exogenous stock to relatively less exogenous stocks, the dominant eigenvector is a natural choice for ordering criteria. Second, we note in this connection that the dominant eigenvector has an alternative interpretation in network theory. It represents eigenvector centrality (in short, EVC) which measures degree of influence of nodes in a graph in a recursive manner. We note that eigenvector centrality also turns out to be a limiting case of Katz-Bonacich centrality as we explain later in the paper, making eigenvector centrality a natural candidate for the analysis. Therefore, in a nutshell, eigenvector centrality allows us to create a unique hierarchy of stocks within the maximally connected component, which also provides an ordering criteria for estimation of unique impulse response functions obtained from VAR.

In principle, one could also use the eigenvector centrality of the volatility correlation matrix as an ordering criteria for identification. However, we see two strong rationale to use eigenvector centrality of the return correlation matrix ( $c_r^{eig}$  henceforth). First, it is fairly stable across time. We divide the total sample which is 16 years long (2002-2017), into four equally non-overlapping wide windows ranging over 4 years viz. 2002-05, 2006-09, 2010-13 and 2014-17 and observe the evolution of  $c_r^{eig}$  in these four time window (table 1). It is evident that the correlation values of pairs of  $c_r^{eig}$  across time are greater than 0.5 in 5 out of 6 cases. We also note that the corresponding values for  $c_v^{eig}$  is considerably low (table 2). Second, definition and measurement of return is unique whereas volatility can be measured in multiple ways. More importantly, the volatility estimates obtained from different methods depend on the specification of the structure of the underlying model. For example, as opposed to the simple GARCH structure, Engle and Figlewski (2014) presents a different model of dynamic linkages in implied volatility (EGARCH/DCC) which is less suitable for modeling large number of assets. Hence, identification criteria based on ordering of volatility centrality would be sensitive to model specification and hence, non-robust. In comparison,  $c_r^{eig}$  is uniquely defined and hence, more credible as an identification criteria.

We represent propagation of the volatility shocks over time in the network using orthogonalized impulse response function of the estimated VAR model. Impulse response function allows us to observe the path

and intensity of the shock propagation for any given origin (i.e. shocks originating in any given node of the network).

Finally, we provide analytical reasons for using  $c_r^{eig}$  and also explore the explanatory factors of the  $c_r^{eig}$ . Note that  $c_r^{eig}$  is a purely statistical measure based on asset returns. We use the bounds of Marčenko-Pastur distribution from random matrix theory (Marčenko and Pastur (1967)), to show that  $c_r^{eig}$  corresponds to a statistically significant eigenvalue. Earlier work has noted that for sectoral indices,  $c_r^{eig}$  is strongly correlated with size of the sector measured by market capitalization, revenue as well as employment (Sharma et al. (2017)). In the current paper, we analyze the relationship between  $c_r^{eig}$  and log of assets, return on assets, credit rating and number of analyst covering the stock. These variables represent size effect, profitability, financial health and popularity respectively.  $c_r^{eig}$  shows strong association with size after controlling for return on asset, credit rating and number of analysts (Sec. 4.2). The coefficient sign of the control variable in the panel regression also supports the common wisdom. Profitable firms are in the core of the network and while financial firm with bad ratings are on the periphery.

Our paper is part of a growing literature that analyzes the mechanism of shock propagation on financial networks. However, we note that our methodology can be suitably applied to other networks as well. Recent empirical work in this area has two related but distinct developments. One, a range of methods have been proposed to construct the network from asset prices. Two, once the network has been constructed then the analysis is done to characterize shock propagation. To construct the asset network, two natural candidates are the return series and the volatility series. Some authors have constructed and analyzed the return network. For example, Geraci and Gnabo (2018) proposed a measurement of connections based on a Bayesian vector autoregression model on the return series. There is also a prior tradition of modeling asset networks based on the realized return correlation matrix (e.g. Mantegna (1999), Tumminello et al. (2010), Pozzi et al. (2013), Sharma et al. (2017)). We adopt this approach to construct our asset network. However, since we analyze shock propagation in this paper, we have modeled the evolution of volatility shock diffusion on the network.

An important methodological point is that the volatility is not observed and has to be estimated from the data. Most notably Diebold and Yilmaz in a series of articles (Diebold and Yilmaz (2014), Diebold and Yilmaz (2015b), Demirer et al. (2017)) have developed a method to construct the *implied* volatility network from asset returns using long-range variance decomposition technique in a vector autoregression set up. The network itself characterizes contributions of individual entities to the others' volatilities. Diebold and Yilmaz (2015a) contains a series of applications of the method to different sorts of financial data. However, the tenability of some the assumption made in this approach is questionable. For example, Diebold and Yilmaz (2014) used a generalized variance decomposition which requires normally distributed shocks. Given the distributional characteristics of the stocks, this assumption is not tenable. In their work, they carried out estimation of volatility using the intra-day return data using the formula proposed by Garman and Klass (1980). In the present context, for simplicity and data availability we use a *generalized autoregressive conditional heteroschedasticity* or GARCH model to estimate the volatility process. Other modeling techniques like the one proposed by Engle and Figlewski (2014) uses a EGARCH/DCC model to estimate the time-varying correlations between implied volatilities of stocks. Implied volatilities contain information about expected future volatility as well as risk premia. Since we focus on realized spillover effects across stocks, we maintain that the volatility estimation can be done with past empirical volatility. The other big departure from this literature is that we do not jointly estimate all volatility series as in DCC model. This approach is quite prevalent in the literature on shock propagation (Diebold and Yilmaz (2015a)). However we differ significantly from this literature in terms of network construction. We utilize a *minimum span-*

ning tree filtered from the return correlation-based network rather than using network constructed from long-range variance decomposition. In the methods section (Sec. 2), we discuss the benefits of this choice. Finally, we estimate a many-dimensional vector autoregressive model on the estimated volatility series with identification restrictions implied by the return network. This last step allows us to actually estimate the impulse response functions.

The idea that there are spill over effects in the financial markets is known in the finance literature. Dessaint et al. (2018) presented strong evidence that in presence of noise shock (which is of non-fundamental nature) to the peer firms' stock prices, firms make changes in their investment decisions due to the limited ability of the managers to filter out information component only from the observed stock prices. Foucault and Fresard (2014) quantified that 1% increase in peer valuation in the stock market impacts the corporate investment of the following firms by about 5.9%. Thus financial information obtained through stock prices have real effects. The opposite has also been documented. Cohen and Frazzini (2008) showed that return predictability exists between firms that are economically linked. Here in our paper, we do not explore the connection between stock market to economic fundamentals and vice versa. Instead we focus solely on quantification of the spill overs within the financial time series.

For us the most relevant work is Barigozzi and Hallin (2017a) who have utilized a sparse vector autoregression framework. They used eigenvector centrality obtained from the residuals in the GARCH specification to identify the long run variance decomposition network. Our work differs significantly in terms of identification criteria and the robustness thereof. We utilize the eigenvector centrality obtained from the *ultra-metric* graph of the stock network (Mantegna (1999)) to rank order the stocks as an identification strategy. We show that the rank-ordering criteria can be constructed easily from the observed return correlation matrix upon non-linear transformation. More importantly, we analyze the rank order of stocks in four different time windows and the rank order is quite stable. As an alternative strategy, we considered eigenvector centrality obtained from the conditional volatility series (obtained from GARCH estimation) correlation network as well. The corresponding rank order is found to be quite unstable across different time windows compared to the one obtained from return network. Thus we show that the proposed identification criteria is considerably robust across time and performs better than the one based on volatility estimation.

The rest of the paper is structured as follows. In Sec. 2, we describe the data and the econometric methodology. This section ends with a step-by-step description of the method proposed for implementation purpose. In the next section, we conduct analysis with impulse response functions over the network to characterize the ripple effects. In Sec. 4, we provide justification of the identifying restrictions. Sec. 5 summarizes the paper and concludes.

## 2 Empirical data and methodology

In the following, we provide the complete description to characterize the shock propagation across the asset network. The background of each of the methods used have been described. In Sec. 2.3 we provide a concise set of steps that encapsulates the entire analysis.

## 2.1 Data description

We have collected data on  $N = 100$  stocks with largest market capitalization in the New York Stock Exchange over a period of 16 years (2002-2017).<sup>5</sup> We denote the entire time period by  $T$ . All stocks are chosen such that they have existed throughout the entire time period. We will divide the time period into four equally wide windows ranging over 4 years viz. 2002-05, 2006-09, 2010-13 and 2014-17 (notationally, we use  $T_1$ ,  $T_2$ ,  $T_3$  and  $T_4$ ; see tables 4, 5, 6 and 7 for stock names, ticker symbols and sectoral affiliations). The first window covers the time when the US economy was in a boom, the second window captures the time of financial crisis, the third captures the time period after crash when the economy was still recovering and the fourth captures a period of relatively stability.

## 2.2 Return and volatility series

Each stock price is denoted by  $\{p_{it}\}_{i \in N, t \in T_j}$  where  $j = 1, 2, 3$  and 4. Return series is defined as the first difference of the log price series

$$r_{it} = \log(p_{it}) - \log(p_{i,t-1}) \quad (1)$$

for  $i \in N, t \in T_j$  where  $j = 1, 2, 3, 4$ . Next, we construct conditional volatility series from return series. There is a plethora of methods that can be used for constructing the latent volatility series. We consider a popular and well-known method of GARCH modeling. We note that this step can be done in multiple ways. Notably, Diebold and Yilmaz (2015a) has popularized a formula due to Garman and Klass (1980) that utilizes daily highest prices, lowest prices, opening prices and closing prices, all measured in logs for all stocks. Because of the substantially lower informational requirement and the fact that our fundamental objective is to characterize spillover from one return series to another, we use GARCH methodology because of its simplicity.

**Estimation of conditional volatility:** We also use a parametric approach to model volatility by using a GARCH( $p, q$ ) framework from each of the return series  $\{r_{it}\}$ . We follow the standard definition of the GARCH model that

$$\begin{aligned} r_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= c + \sum_{i=1}^p \alpha_i r_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2. \end{aligned} \quad (2)$$

We will denote conditional volatility of the  $i$ -th stock by  $\sigma_{it}$ .

### 2.2.1 Selection of maximally connected component through Lasso

In order to analyze the mechanism of shock propagation, we first need to identify the set of stocks which can be affected by the shock propagation (Hsu et al. (2008)). Towards that objective, we first construct *maximally connected component*  $G_{mcc}$  of the network  $G$ .

<sup>5</sup>All data (prices as well as other company-level information) have been collected from Thomson Reuters Eikon database.

**Definition 1 (Maximally connected component:)** Given a network  $G = (V, E)$ , we call a subnetwork  $G' \subseteq G$  with structure  $G' = (V', E')$  a maximally connected component of  $G$ , if  $G'$  is connected and for all nodes  $u \in V$  and  $u \notin V'$ , there exists no node  $v \in V'$  such that  $(u, v) \in E$ .

We identify the *maximally connected component*  $G_{mcc}$  by estimating a Lasso (Least Absolute Shrinkage and Selection Operator; Tibshirani (1996)) model. For a general description of Lasso, consider the following least-square estimate with penalty for  $l^1$  norm over the estimated parameters:

$$\min \sum_{t=1}^T \left( y_t - \sum_{j=1}^k \beta_j x_{jt} \right)^2 \quad \text{s.t.} \quad \sum_{j=1}^k |\beta_j| \leq C \quad (3)$$

where  $y_t$  represents a vector of response variables and  $x_{jt}$  represents explanatory variables where  $j \in k$ . Essentially, we fit the sparse VAR model in the form of Lasso regression described above with the obvious substitution of  $\{x_{jt}\}$  by lagged values of  $y_{jt}$ . For actual implementation, we consider the sequence of conditional volatility  $\{\sigma_{it}\}_{i \in N, t \in T_j}$  estimated above through GARCH procedure (where  $j \in \{1, 2, 3, 4\}$ ) and utilize adaptive Lasso technique (Zou (2006)) which optimally estimates the model and has the ability to robustly identify the right subset model with true predictor variables (oracle properties). After we fit the model (allowing one lag), a substantial number of the links between all  $\{i, j\}$  pairs would be set equal to zero. Let us denote estimated coefficient matrix by  $\beta$ . Then we convert into a binary matrix (0, 1) where all nonzero elements are converted into ones and the zeros remain zero. Note that the row sum and the column sum of this binary transformation of the  $\beta$  matrix represents *indegree* and *outdegree* respectively.<sup>6</sup> Then we use a threshold to differentiate between connected and unconnected nodes. For the empirical exercise, we followed the convention that if a node has both *indegree* and *outdegree* less than 10%, then it does not belong to the maximally connected component  $G_{mcc}$ . Thus in order to belong to the connected component, one node needs to be connected with at least 10% nodes with either incoming or outgoing links, or both. If the threshold is set at 0, then we simply recover the original set of stocks. Notationally, let us assume that after selecting the  $G_{mcc}$ , the number of surviving stock is  $n \leq N$ .

### 2.2.2 Filtration of edges to construct a hierarchical network

Once the connected component has been extracted, we carry out all operations on that that set of stocks only. We construct a sample correlation matrix  $\Gamma^\phi$  from both the return series  $\{r_{it}\}$  and the volatility series  $\{\sigma_{it}\}$ , by using pairwise correlation coefficients defined as

$$\gamma_{ij}^\phi = \frac{E((\phi_i - \bar{\phi}_i)(\phi_j - \bar{\phi}_j))}{\sigma_i^\phi \cdot \sigma_j^\phi} \quad (4)$$

for  $i, j \in n$ , where  $\bar{x}$  denotes mean value of  $x$  and  $\sigma^\phi$  denotes unconditional volatility of  $\phi \in \{r, \sigma\}$ .

We proceed to construct two hierarchical networks (Di Matteo et al. (2010)) based on the two correlation matrices ( $\Gamma^r$  and  $\Gamma^\sigma$ ). We note that individual correlation values cannot serve the purpose of distance as they can be negative. Therefore, following the literature (Mantegna (1999), Bonanno et al. (2004)) we consider

<sup>6</sup>*Indegree* represents how much a particular node in a graph is open to influence of other nodes and *outdegree* represents direct influence of one particular node on the remaining nodes.

the metric (Gower (1966))

$$d_{ij} = \sqrt{2(1 - \rho_{ij})}. \quad (5)$$

Using this nonlinear transformation, we can construct two distance matrices  $D^r$  and  $D^\sigma$  created respectively from the two correlation matrices  $\Gamma^r$  and  $\Gamma^\sigma$ . With the help of the distance matrix, we construct two networks  $G^r = (V^r, E^r)$  and  $G^\sigma = (V^\sigma, E^\sigma)$ , where  $V$  denotes the set of nodes/vertices (stocks) and  $E$  denotes the set of edges obtained from Eqn. 5. Clearly these two networks contain  $n$  nodes and all the edges are of positive magnitude.

In order to filter out the minimal sized network that has the maximum information (Bonanno et al. (2004)), we construct the *minimum spanning tree* (MST hereafter) from  $\Gamma^r$  and  $\Gamma^\sigma$ . For the purpose of visualization, we only use the MST obtained from  $\Gamma^r$ .<sup>7</sup> Below we define MST.<sup>8</sup>

**Definition 2 (Minimum spanning tree:)** *A spanning tree of a graph  $G$  is a connected subgraph of  $G$  that has no cycles and connects all vertices with minimum number of edges. For a weighted graph, the minimum spanning tree is a spanning tree that has the least weight among all possible spanning trees.*

Thus MST identifies the strongest links across the graph  $G_{mcc}$ .<sup>9</sup>

To summarize the development so far, we have identified the connected component of stocks (Sec. 2.2.1) and constructed the minimum spanning tree which preserves the links across stocks that are the strongest under aggregation. We will utilize the MST to graphical exposition of the ripple effect as it drastically reduces the number of links or edges to  $n - 1$ .

### 2.2.3 Identification restrictions from network topology

Once, we have constructed the maximally connected component of the network, we proceed towards quantifying the spillover effects through estimation of a vector autoregression model. First, we need to specify the identification restrictions.

We identify the VAR model through centrality measures. We consider two criteria: Eigenvector centrality and Katz-Bonacich centrality (Bonacich (1987)) which are defined as follows.

**Definition 3 (Eigenvector centrality:)** *For a given network  $G = (V, E)$  with adjacency matrix  $A$ , we define eigenvector centrality to be a vector  $c^{eig}$  which solves*

$$Ac^{eig} = \lambda c^{eig} \quad (6)$$

where  $\lambda$  is chosen to be the maximum eigenvalue  $\lambda_{max}$  of the adjacency matrix  $A$ .

<sup>7</sup>As we will show later that a robust identification restrictions can be obtained from  $\Gamma^r$ , which we utilize for estimation purpose.

<sup>8</sup>There are many standard algorithms to create *MST* from a fully connected graph. Prim's and Kruskal's algorithms are standard ones. Also, *MST* is unique for a graph when all edges have distinct weights (which is the case for an adjacency matrix obtained from correlation matrix).

<sup>9</sup>Note that due to the transformation of the correlation matrix (Eqn. 5), minimum weight in the distance matrix refers to maximum correlation. Therefore, *MST* on the distance matrix identifies the strongest links in terms of correlations.

**Definition 4 (Katz-Bonacich centrality:)** For a given network  $G = (V, E)$  with adjacency matrix  $A$ , we define Kat-Bonacich centrality to be a vector  $c^{KB}$  such that

$$c^{KB}(\alpha) = ((I - \alpha A^T)^{-1} - I) \vec{1}, \quad (7)$$

where  $\alpha$  is an attenuation factor (imposed from outside),  $I$  is an identity matrix and  $\vec{1}$  is a vector of ones. Under the condition that  $\alpha < 1/\lambda_{max}$  ( $\lambda_{max}$  being the highest eigenvalue of  $A$ ), the centrality measure is well-defined.

For simplification of exposition of the empirical analysis, we consider the unique ranking produced by eigenvector centrality  $c^e$  which is known to be the limit of the ranking produced by Katz-Bonacich centrality as stated in Thm. 1 below.

**Theorem 1 (Theorem. 3.3 from Benzi and Klymko (2013))** As  $\alpha \rightarrow \frac{1}{\lambda_{max}}$  (i.e. converges from below), the ranking produced by  $c^{KB}(\alpha)$  converge to that produced by  $c^{eig}$ . When  $\alpha \rightarrow 0$ , then ranking produced by  $c^{KB}(\alpha)$  converge to that produced by degree of nodes.

Although it is quite feasible, we have not reported the results for  $\alpha \rightarrow 0$  as shock propagation in a financial network depends on the nature of the influential nodes rather than number of connections.

The relative exogeneity of the stocks can be defined as having highest centrality in the correlation matrices  $\Gamma^r$  and  $\Gamma^\sigma$ . In this paper, we will argue that  $\Gamma^r$  is a substantially better exogeneity criteria due to its stability over all time periods (Sec. 4). This relative exogeneity criteria would be used in the next step for ordering of the stocks through Cholesky decomposition. Thus the estimation procedure for the vector autoregression model would depend on recursive identification.

We note that one can also utilize the centrality measures defined over the distance matrices  $D^r$  and  $D^\sigma$ . Clearly, because of the formula we are using for distance (Eqn. 5), high centrality in  $\Gamma$  would be reflected as low centrality in  $D$ . Empirically, the correlation is around -0.99. Hence, either way we get almost identical ordering.

#### 2.2.4 Vector autoregression estimation and orthogonalization

In Sec. 2.2, we have constructed the volatility series for  $n$  number of stocks. By applying the technique presented in Sec. 2.2.1, we can find the maximally connected component of the stock network ( $G_{mcc}$ ) created from the empirical correlation matrix (Sec. 2.2.2). Let us write the corresponding structure as  $G_{mcc} = (V_{mcc}, E_{mcc})$ . Clearly, the number of the surviving nodes in the resultant connected component is lesser than the original network i.e.,  $n_{mcc} \leq n$ .

Now, we select the corresponding volatility series  $\{\sigma_{it}\}_{i \in n_{mcc}}$ . We estimate a vector autoregression model with  $p$ -lags of the following form<sup>10</sup>:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + u_t \quad (8)$$

<sup>10</sup>This is a generic description of an VAR model. For actual estimation, we substitute  $y_{it}$  by  $\sigma_{it}$

where all matrices  $\{A_i\}_{i=1,\dots,p}$  have dimensions of  $n \times n$ . The noise term has a covariance matrix  $E(u_t u_t^T) = \Xi$ .

Generally, for lag selection one would rely on information criteria. We find that apart from AIC, all other information criteria chooses a lag of 1. Also, choosing lag 1 makes the model much more economized in terms of parameters, noting that the number of the variables is large ( $n \sim 100$ ) and the time horizon is about 1000 days (about 250 working days per year). Finally, more lags would not necessarily provide any economic rationale for shock propagation. Therefore, all the results reported here are obtained by allowing only one lag. Thus we consider the above-mentioned VAR to represent a simplest but non-trivial system capturing the dynamics of volatility across stocks.

To construct orthogonalized impulse response functions, we first estimate the covariance matrix  $\hat{\Xi}$  of the error term  $u$ . Following standard procedure advocated by Sims, we conduct a Cholesky decomposition of the matrix  $\hat{\Xi}$ . Let us denote the resultant lower-triangular matrix  $\xi$ , i.e.,

$$\hat{\Xi} = \xi \xi^* \quad (9)$$

where  $\xi^*$  denotes the conjugate transpose of  $\xi$ . Then we pre-multiply Eqn. 8 to find the orthogonalized impulse response functions.

### 2.2.5 Ripple effects through impulse response functions

After estimating the model, we compute the impulse response functions with the estimated coefficient matrices. Given any shock in any of the stock volatility series, we can compute the corresponding impulse responses. We show an example in Fig. 1. To enhance visual clarity, we have utilized the minimum spanning tree and depicted the propagation of volatility shock emanating from a chosen epicenter over four different time points.

## 2.3 Algorithm to characterize the ripple effects

Here, we provide a step-by-step algorithm to construct the asset network and characterize the shock propagation mechanism.

1. Choose  $N$  number of stocks to create return series over  $T$  time periods (e.g. days) such that  $T \gg N$ .
2. Estimate volatility series from each of the return series by using  $GARCH(p, q)$  model.
3. Estimate a Lasso model to find maximally connected component  $G_{mcc}$  from the network of stocks  $G$ .
4. Construct a network based on the correlation structure of the surviving stocks  $G^r$ . For visualization, extract the minimum spanning tree.
5. Estimate a vector autoregression model on the volatility series from all stocks in  $G_{mcc}$ . The model is identified by imposing an ordering obtained from the eigenvector centrality  $C_r^{eig}$  of the return correlation matrix.
6. Characterize the ripple over the whole network by using estimated impulse response functions.

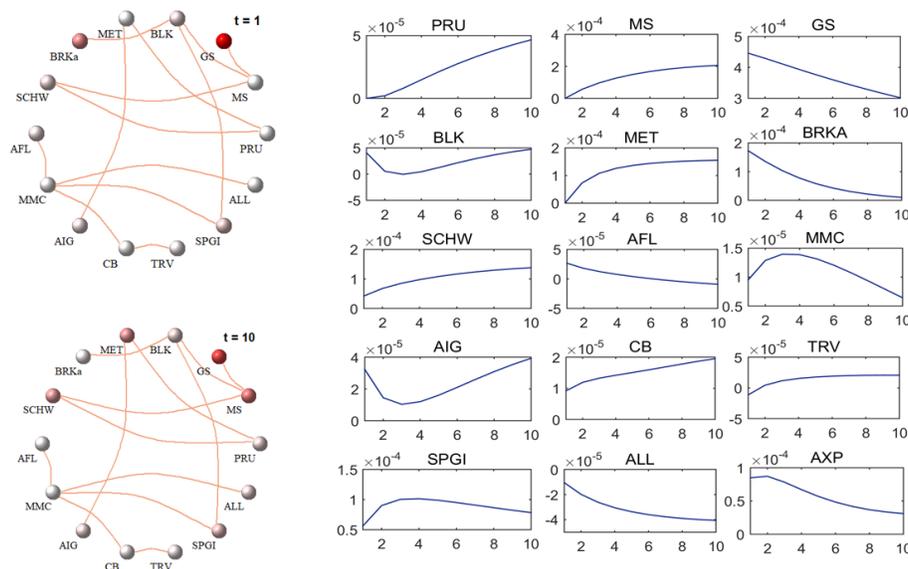


Figure 1: (color online) Ripple effects on the financial network during the period 2014-17. We have chosen all financial firms within the set of 100 firms with largest market capitalization in the NYSE, to construct the network. The diagram depicts propagation of a positive volatility shock given to Goldman Sachs (GS) across the network (for visualization we have utilized a minimum spanning tree) at two time points  $t = 1$  (left-top panel) and 10 (left-bottom panel). Intensity of color represents the magnitude of return volatility of the nodes in the network. Right panel: Impulse response functions resulting from the same shock.

In the following, we utilize the algorithm to characterize the *ripples* across the financial networks, emanating from chosen epicenters. we also characterize how the impact of the ripples reduces over time.

### 3 Empirical analysis

We implement our methodology on multiple sets of data. First, we analyze intra-sectoral spill over effects. Second we apply to a large number of stocks listed in the New York stock exchange with the highest market capitalization. As we will demonstrate below, it is easier to visualize the ripple effects of a volatility shock over networks rather than the standard time series view of impulse response functions. First, we demonstrate the mapping between the impulse response functions and the network diagrams. Then we exclusively utilize the networks to display the ripple effects. For visualization purpose, we use a color coding on the nodes of the network, where magnitude of an impact is captured through density of color.

### 3.1 Intra-sectoral spill over effects

Although the methodology can be utilized to model the spill over effects across stocks in any sector, we have chosen to elucidate the mechanism with financial and industrial sectors. We first consider the financial sector during the period 2014-17. In particular, we analyze the results of a positive volatility shock given to Goldman Sachs Group Inc. (GS).<sup>11</sup>

Fig. 1 shows the responses of an unit orthogonalized volatility shock emanating from Goldman Sachs Group Inc. (GS) to all other financial stocks (which belong to the set of top 100 stocks in New York stock exchange in terms of market capitalization). On the left, we show the network of stocks at  $t = 1$  (time when the volatility shock hits) and the same at  $t = 10$ . The color intensity denotes the strength of the shock spill over across nodes of the network. We see that Berkshire Hathaway Inc. (BRKA) was immediately affected and recovered by period 10. But Morgan Stanley (MS) remains affected even after 10 periods, followed by Charles Scwab Corp (SCHW) and MetLife Inc. (MET). The right panels shows complete descriptions of the impulse response functions across periods  $t$  ranging from 1 to 10.

Next, we consider the effects of a positive volatility shock given to Honeywell International Inc. (HON). Fig. 2 shows the propagation of the shocks to all other industrial stocks (belonging to the set of top 100 stocks with largest market capitalization in the New York stock exchange during 2014-17) over four periods:  $t = 1$ ,  $t = 4$ ,  $t = 7$  and  $t = 10$ . As the results exhibit, the effect on other stocks are not substantial after 10 periods.

### 3.2 Economy-wide spill over effects

Above, we have considered only within sector shock propagation. In this section, we apply the proposed methodology to the full data set of 100 stocks with largest market capitalization. We conduct two exercises to model the ripple effects.

One, we split the data in four samples covering non-overlapping periods 2002-05, 2006-09, 2010-13 and 2014-17. In each of them we apply a positive volatility shock to Goldman Sachs Group Inc. (GS) and trace out the shock propagation across the network. We show that there is considerable variations in the spill over effects across different time periods. Second, we quantify the above finding that shock spill overs can be different across different periods, by combining the information of shocks to all firms across all periods. We show that generally some periods are more prone to shock spill over than others. Below, we describe the specific results and provide the corresponding details.

Fig. 3 shows the ripples emanating from Goldman Sachs Group Inc. (GS) in response to a positive volatility shock in the period 2014-17. We have followed the entire procedure described in Sec. 2.3. We have plotted the minimum spanning tree on the maximally connected component selected by Lasso estimation. The diagram shows four snapshots at time  $t = 1, 4, 7$  and  $10$ . We see that firms like Morgan Stanley (MS) and Charles Scwab Corp (SCHW) among others are affected. Note that it is not very obvious that Goldman Sachs and Morgan Stanley would share a direct connection. Since minimum spanning tree chooses a tree from all possible connections across all possible pairs of stocks, there is no guarantee that firms belonging to the same sector will be placed close to each other on the minimum spanning tree. However, this tends to

<sup>11</sup>Complete description of the company names and the corresponding short forms (ticker symbols) can be found in the appendix. See tables 4, 5, 6 and 7 for listed companies during the periods 2002-05, 2006-09, 2010-13, 2014-17 respectively.

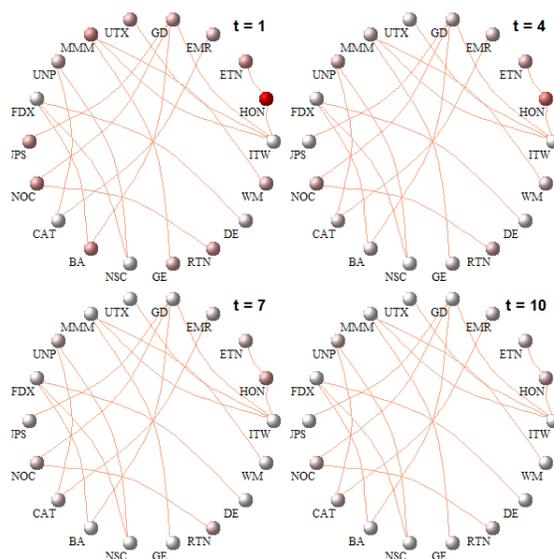


Figure 2: (color online) Ripple effects on the industrial network during the period 2014-17. We have chosen all industrial firms within the set of 100 firms with largest market capitalization in the NYSE, to construct the network. The diagram depicts propagation of a positive volatility shock given to Honeywell International Inc. (HON) across the network at four time points  $t = 1$  (left-top panel), 4 (right-top panel), 7 (left-bottom panel) and 10 (right-bottom panel). Intensity of color represents the magnitude of return volatility of the nodes in the network.

happen when pairs of stocks comove a lot. Responses to the same shock emanating from the same epicenter Goldman Sachs Group Inc. (GS) have been shown in the Appendix (Fig. 5, 6, 7). Interestingly, as the figures suggest the effects were much stronger during the period 2002-05. After that the effects subdued substantially during two consecutive periods 2006-09 and 2010-13.

Analysis of the effects of only one stock does not give a complete picture of the market as a whole. But analyzing all separate responses for all stocks would be difficult. Therefore, we summarize the information of the impact of individual stocks on the remaining market in Fig. 4. On the  $x$ -axis, we show four periods 2002-05, 2006-09, 2010-13 and 2014-17. On the  $y$ -axis, we show the stock identities in terms of numbers (can be mapped to the lists given in the Appendix; viz. tables 4, 5, 6 and 7) which act as the epicenters of the ripple effects. In successive experiments, we give unit shocks to the individual epicenters and trace the ripples emanating from the epicenters spreading throughout the stock network. The  $z$ -axis shows the number of firms showing response of at least 5% magnitude of that of the epicenter after 4 time points i.e.,  $t = 4$ . We see that during 2002-05, the average response ( $\sim 30$ ) was lesser compared to other periods. The next period 2006-09 showed more pronounced responses ( $\sim 36$ ) with 2010-13 showing the highest average number of stocks responding ( $\sim 40$ ). The last period 2014-17 again comes back closer to the first period ( $\sim 33$ ).



## 4 Return centrality as identification criteria

In this section, we discuss the economic reasons for considering return centrality as the identification criteria. The main motivation comes from eigenvalue decomposition of cross-correlation matrix of return series. Let us denote the cross-correlation matrix by  $\Sigma_r$ . We can decompose the matrix as

$$\Sigma_r = \sum_{i=1}^n \lambda_i c_i c_i^T \quad (10)$$

where  $\lambda_i$  is the  $i$ -th eigenvalue and  $c_i$  is the  $i$ -th eigenvector. Let us denote the highest eigenvalue in modulus is  $\lambda_1$  and the corresponding eigenvector is  $c_1$  (and without loss of generalization,  $\lambda_i \geq \lambda_j$  for  $i < j$ ). Interestingly,  $c_1$  has two interpretations. From a graph theoretic point of view, it represents the eigenvector centrality. This is precisely what we are using for identification (we labeled it  $c_r^{eig}$ ). On the other hand, from a financial point of view it represents the relative weights of different stocks in a market portfolio. Thus it represents the relative contribution of different stocks to the market-wide movement (outgoing) or the impact of exogenous shocks on different stocks (incoming).

We need to make sure that the dominant eigenvalue  $\lambda_1$  is significantly large and is not an artifact arising out of random noise. For that purpose, we utilize a well known result from *random matrix theory* that provides an upper bound to the spectral radius of a correlation matrix generated from  $N$  random time series of length  $T$ . Such matrices are known as *Wishart* matrices.

**Theorem 2 (from Marčenko and Pastur (1967))** *Let  $N \rightarrow \infty$  and  $T \rightarrow \infty$  with  $Q \equiv N/T > 1$ . Consider a Wishart matrix  $W = XX'$  where  $X \sim N(0, I)$ . The upper bound on the modulus of the maximum eigenvalue of  $W$  is given by*

$$\lambda_{u.b.} = \left(1 + \frac{1}{\sqrt{Q}}\right)^2. \quad (11)$$

We verified it empirically that  $\lambda_{u.b.} < \lambda_1$ , the maximum eigenvalue.

### 4.1 Centrality measures: Stability over volatile and tranquil periods

Here, we have to make choice for the identification criteria. We note that both of the dominant eigenvectors of the return correlation matrix and the volatility correlation matrix, could have been candidates for identification criteria in terms of relative exogeneity. We note that return is an observed variable and hence, free of measurement noise and/or model specification errors. Volatility on the other hand, is latent and hence needs to be estimated. Such estimation results will be model specific. Thus return centrality is a more robust choice.

More importantly, we find that the centrality measure based on the return matrix is considerably more robust and stable across different timing horizon. We show the correlation results of the firms' return centralities in table 1 in four different snapshots (2002-05, 2006-09, 2010-13 and 2014-17) that covers pre-crisis to post-crisis periods. In table 2, we show the same for volatility correlation centrality. Return centralities are clearly much more correlated across time than the corresponding volatility centralities. Thus the return centrality measure is more robust.

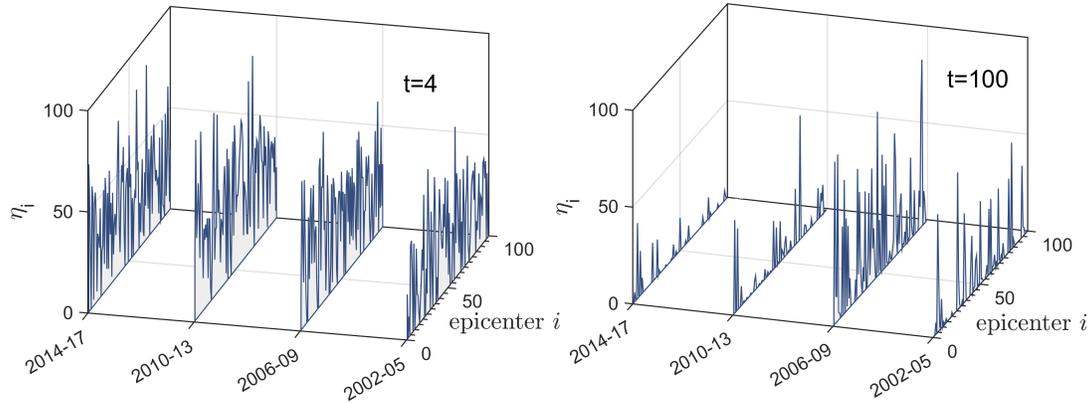


Figure 4: (color online) A summary of impulse response functions ( $z$ -axis) across all stocks ( $y$ -axis) across all four periods ( $x$ -axis). On the  $z$ -axis we plot the number of stocks showing at least 5% response (in absolute values) with respect to a shock given to the epicenter at  $t = 1$  (listed on the  $y$ -axis, from 1 to 100) after 3 time points i.e., at  $t = 4$ . Intensity of color represents the magnitude of return volatility of the nodes in the network.

Table 1: Correlation matrix: Return Centrality  $c_r^{eig}$

	$c_r^{eig}$ (02-05)	$c_r^{eig}$ (06-09)	$c_r^{eig}$ (10-13)	$c_r^{eig}$ (14-17)
$c_r^{eig}$ (02-05)	1	–	–	–
$c_r^{eig}$ (06-09)	0.641	1	–	–
$c_r^{eig}$ (10-13)	0.508	0.564	1	–
$c_r^{eig}$ (14-17)	0.511	0.405	0.685	1

Table 2: Correlation matrix: Volatility centrality  $c_v^{eig}$

	$c_v^{eig}$ (02-05)	$c_v^{eig}$ (06-09)	$c_v^{eig}$ (10-13)	$c_v^{eig}$ (14-17)
$c_v^{eig}$ (02-05)	1	–	–	–
$c_v^{eig}$ (06-09)	0.375	1	–	–
$c_v^{eig}$ (10-13)	0.240	0.475	1	–
$c_v^{eig}$ (14-17)	0.088	0.294	0.442	1

## 4.2 Explanatory factors of return centrality $c_r^{eig}$

So far we have utilized the return centrality  $c_r^{eig}$  as an identification criterion. Fundamentally, it is a measure which has extracted from return dynamics of different stocks. In this section, we explore the its relationship with firm-level fundamentals. Since the variable of interest here is a centrality measure, we create a simple regression specification where centrality (from the return correlation matrix) is a function of firm size (log of the total asset), return on assets (ROA; as a proxy of profitability), financial rating (as a proxy for health of the firm; higher rating indicates worse health) and number of analysts (as a measure of profitability). Here we note that we are unaware of any literature that relates firm-level economic variables to centrality of the firms in the in return/volatility network (except Sharma et al. (2017) which shows that economic size and return centrality are correlated at the sector level). Therefore, we analyze only the known and well-accepted factors that explain variation in expected return. We find that size, profitability, financial health and popularity explain variation in centrality as described below. However, we do not expect our list of such variables to be exhaustive.

Harvey et al. (2016) provides a comprehensive list of potential variables that explain expected return of firms. Here, we choose a set of four firm-level variables that have been documented to have explanatory power for expected returns. Banz (1981) demonstrated that higher market capitalization is correlated with lower expected return. This observation had been incorporated in the three-factor model propounded by Fama and French (1992), where size is shown to be a systematic risk factor. In this context, we note that Sharma et al. (2017) showed that sectors with higher size (proxied by market capitalization, total revenue and employment) tends to have higher centrality in the sectoral network constructed from the observed return correlation matrix. Here we proxy size by log of assets, but our results hold true for log of market capitalization as well. We find that larger firms typically are more central in the financial network. Next two variables we consider represent profitability and financial health of the firm. Novy-Marx (2013) showed profitability measured by gross return of assets at the firm level predicts variation in expected return. Also, financial health explains variations in expected returns and is shown to correlate with size effect and the value effect on return (Vassalou and Xing (2004)). Therefore, as a summary measure of financial health, we incorporate credit rating of firms. Finally, we test the effects of popularity of stocks by considering number of analysts following a stock. Hameed et al. (2015) documented that analysts have differential propensity to follow firms based on observed correlation of the firm-level fundamentals with the corresponding industry peers. They showed that firms with higher analyst coverage impacts stock prices of firms with lower coverage, when analysts revise earning forecasts of those firms. But the opposite does not happen. This unidirectional spillover of information provides a rationale for higher co-movements between stocks with larger analyst coverage, which will manifest in the observed return correlation matrix.

Table 3 gives the result of these regression for all four time windows (2002-06, 2006-09, 2010-13 and 2014-17). Log asset is generally strongly correlated with the centrality measure with negative coefficients.<sup>12</sup> An immediate implication is that shock propagates from large firm to smaller firms. Profitability (measured by ROA), financial health (measured by ranking based on credit rating where higher ranking implies lower financial health) and analysts' coverage does not seem to have consistent explanatory power for all snapshots.

<sup>12</sup>Note that due to transformation via Eqn. 5, a stock belonging to the core with high centrality in the distance matrix will have relatively weaker correlations with other stocks. Thus higher firm size is associated with firms at the periphery of the network obtained from distance matrix. Therefore, the firms with higher presence in the market mode of the asset returns (Plerou et al. (2002)) are empirically found to be typically larger in size than the ones with smaller presence.

Finally, we estimate panel regression with all the firms (those who were operational from 2002 to 2017). A random effect model (preferred by Hausman test) with time effect retains the feature that core firms (in the distance matrix) are smaller in size, less profitable and are in worse financial health. Combining with the earlier interpretation, we see that firms that have more presence in the market mode (obtained from the return correlation matrix) are larger in size, more profitable and have better financial health. Popularity does not seem to play a consistent role either way.

## 5 Summary and conclusion

Propagation of shocks on a financial network is a very intuitive phenomenon. The recent literature on financial econometrics has started formally addressing and modeling it. However, the standard approach is to analyze the long-range variance decomposition. Essentially the method creates a network across financial entities (e.g. stocks) by attributing the fraction of variations explained by one entity of the others. In this paper, we propose a complementary approach by estimating a VAR model on the latent volatility processes of the stocks with an identification criteria obtained from the topology of the network. In particular, we show that the centrality measure based on dominant eigenvector of the observed return correlation matrix provides a intuitive and robust identification strategy. The identified model allows us to observe the ripple effects originating from a particular epicenter in the stock network and to quantify the corresponding magnitude of fluctuations.

This paper relates most notably to the work of Diebold and Yilmaz (Diebold and Yilmaz (2015a)) who constructed the above-mentioned method of long range variance decomposition based on VAR models to construct the network across assets. Our proposed approach is to infer the network form the observed cross-correlation data and utilize a conversion to a norm to create a minimum spanning tree, which preserves the strongest links. For the construction of the minimum spanning tree from the stock network, we rely on the work by Mantegna (1999) and others (see for example Tumminello et al. (2010) for a detailed treatment). We take correlation matrix as given and conduct a spectral analysis to find the centrality of the stocks. It is important to note that the centrality measures (in the sense of eigenvector centrality) obtained from the correlation matrix and the distance matrix are almost exactly anti-correlated, which is consistent with the algebraic transformation used for creating the metric from the correlation values. Hence, either ordering can be used for identification purpose by maintaining the appropriate signs.

The centrality measure serves multiple purpose and can be used as a very useful tool. One, it provides a natural interpretation of the core-periphery structure of the network. One can in principle decompose the network into a core and periphery, but centrality provides a more continuous version of relative coreness than a binary characterization of being either in core or periphery. This becomes very useful to define relative exogeneity of the stocks, which acts as an identification criteria for finding orthogonalized impulse response functions through standard Cholesky decomposition. Third, we also utilize a result from random matrix theory (in particular, Marchenko-Pastur theorem developed in Marčenko and Pastur (1967)) to establish the statistical validity of the centrality measure that the result is not outcome of random noise. This approach was used by Plerou et al. (2002) and have become useful for analyzing financial risk with high-dimensional data (see e.g. Bouchaud and Potters (2003)).

Modeling spillover through linkages of volatilities have been an important area of work in the academic as well as policy domain, as it provides a more comprehensive view of the risk embedded in a financial system.

Table 3: Cross section (OLS) and panel (random effect) regression results

	Dependent variable:									
	OLS					panel: Random linear				
	2002-05	2006-09	2009-13	2014-17	(Individual)	(Time)	(Two-ways)			
logasset	-0.284*** (0.076)	-0.065 (0.100)	-0.228*** (0.082)	-0.192** (0.085)	-0.221*** (0.084)	-0.207*** (0.058)	-0.204*** (0.043)	-0.207*** (0.086)		
ROA	-0.003 (0.017)	-0.015 (0.012)	-0.012 (0.018)	-0.012 (0.016)	-0.012 (0.016)	-0.016** (0.007)	-0.015* (0.008)	-0.016 (0.011)		
ratingr	0.001 (0.005)	0.006 (0.005)	0.009*** (0.003)	0.003 (0.004)	0.003 (0.004)	0.004** (0.002)	0.005** (0.002)	0.004 (0.003)		
Analyst	0.005 (0.014)	-0.041* (0.023)	-0.0004 (0.013)	0.024** (0.011)	0.024** (0.011)	0.004 (0.007)	0.012 (0.007)	0.012* (0.011)		
Constant	6.839*** (1.854)	1.590 (2.441)	5.618*** (2.033)	4.767** (2.120)	4.943** (2.225)	4.855*** (1.401)	4.944*** (1.065)	4.856** (2.089)		
Observations	100	98	100	100	100	392	392	392		
R <sup>2</sup>	0.138	0.006	0.066	0.044	0.094	0.049	0.074	0.069		
Adjusted R <sup>2</sup>	0.129	-0.004	0.057	0.034	0.056	0.039	0.064	0.059		
Res.Std.Error	0.933	1.002	0.971	0.983	0.972	0.983	0.972	0.972		
F Statistic	15.727***	3.097**	6.978***	4.526**	2.467*	5.009***	7.679***	7.158***		
df	1; 98	4; 91	1; 98	4; 93	1; 98	4; 387	4; 387	4; 387		

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Diebold and Yilmaz (2008) showed that there is a link between macroeconomic volatilities and financial volatilities. Recent work by Hüser et al. (2018), Corsi et al. (2018), Hué et al. (2019) among others, highlight the role of networks in measurement and contribution to systemic risk. Our paper contributes to the same literature and the proposed algorithmic approach might provide a visual laboratory to analyze risk.

## References

- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origin of aggregate fluctuations. *Econometrica*, 80:1977–2016.
- Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608.
- Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2016). Networks, shocks, and systemic risk. *Handbook of Network Economics*. Edited by Yann Bramoullé, Andrea Galeotti, and Brian Rogers., Oxford University Press.
- Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3–18.
- Barigozzi, M. and Hallin, M. (2017a). Generalized dynamic factor models and volatilities: estimation and forecasting. *Journal of Econometrics*, 201(2):307–321.
- Barigozzi, M. and Hallin, M. (2017b). A network analysis of the volatility of high dimensional financial series. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 66(3):581–605.
- Benzi, M. and Klymko, C. (2013). A matrix analysis of different centrality measures. *arXiv preprint arXiv:1312.6722*, page 37.
- Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology*, 92(5):1170–1182.
- Bonanno, G., Caldarelli, G., Lillo, F., Miccichè, S., Vandewalle, N., and Mantegna, R. N. (2004). Networks of equities in financial markets. *European Physical Journal B*, 38:363–371.
- Bouchaud, J.-P. and Potters, M. (2003). *Theory of financial risk and derivative pricing: from statistical physics to risk management*. Cambridge university press.
- Cohen, L. and Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance*, 63(4):1977–2011.
- Corsi, F., Lillo, F., Pirino, D., and Trapin, L. (2018). Measuring the propagation of financial distress with granger-causality tail risk networks. *Journal of Financial Stability*, 38:18–36.
- Demirer, M., Diebold, F. X., Liu, L., and Yilmaz, K. (2017). Estimating global bank network connectedness. *Journal of Applied Econometrics*.
- Dessaint, O., Foucault, T., Frésard, L., and Matray, A. (2018). Noisy stock prices and corporate investment. *Review of Financial Studies (forthcoming)*.
- Di Matteo, T., Pozzi, F., and Aste, T. (2010). The use of dynamical networks to detect the hierarchical organization of financial market sectors. *The European Physical Journal B*, 73(1):3–11.
- Diebold, F. X. and Yilmaz, K. (2008). Macroeconomic volatility and stock market volatility, worldwide. Technical report, National Bureau of Economic Research.

- Diebold, F. X. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.
- Diebold, F. X. and Yilmaz, K. (2015a). *Financial and Macroeconomic Connectedness: A Network Approach to Measurement and Monitoring*. Oxford University Press, USA.
- Diebold, F. X. and Yilmaz, K. (2015b). Trans-atlantic equity volatility connectedness: Us and european financial institutions, 2004–2014. *Journal of Financial Econometrics*, 14(1):81–127.
- Engle, R. and Figlewski, S. (2014). Modeling the dynamics of correlations among implied volatilities. *Review of Finance*, 19(3):991–1018.
- Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465.
- Foucault, T. and Fresard, L. (2014). Learning from peers' stock prices and corporate investment. *Journal of Financial Economics*, 111(3):554–577.
- Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 53:67–78.
- Geraci, M. V. and Gnabo, J.-Y. (2018). Measuring interconnectedness between financial institutions with bayesian time-varying vector autoregressions. *Journal of Financial and Quantitative Analysis*, 53(3):1371–1390.
- Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53:325–338.
- Hameed, A., Morck, R., Shen, J., and Yeung, B. (2015). Information, analysts, and stock return comovement. *The Review of Financial Studies*, 28(11):3153–3187.
- Harvey, C. R., Liu, Y., and Zhu, H. (2016). ... and the cross-section of expected returns. *The Review of Financial Studies*, 29(1):5–68.
- Hsu, N.-J., Hung, H.-L., and Chang, Y.-M. (2008). Subset selection for vector autoregressive processes using Lasso. *Computational Statistics & Data Analysis*, 52(7):3645–3657.
- Hué, S., Lucotte, Y., and Tokpavi, S. (2019). Measuring network systemic risk contributions: A leave-one-out approach. *Journal of Economic Dynamics and Control*, 100:86–114.
- Hüser, A.-C., Hałaj, G., Kok, C., Perales, C., and van der Kraaij, A. (2018). The systemic implications of bail-in: a multi-layered network approach. *Journal of Financial Stability*, 38:81–97.
- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems*, 11(1):193–197.
- Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457.
- Mullaly, A. R. (2008). Examining the state of the domestic automobile industry.
- Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28.

- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T., and Stanley, H. E. (2002). Random matrix approach to cross correlations in financial data. *Physical Review E*, 65:066126.
- Plosser, C. I. (2009). Redesigning financial system regulation.
- Pozzi, F., Di Matteo, T., and Aste, T. (2013). Spread of risk across financial markets: better to invest in the peripheries. *Scientific Reports*, 3:1665.
- Sharma, K., Gopalakrishnan, B., Chakrabarti, A. S., and Chakraborti, A. (2017). Financial fluctuations anchored to economic fundamentals: A mesoscopic network approach. *Scientific reports*, 7(1):8055.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288.
- Tumminello, M., Di Matteo, T., Aste, T., and Mantegna, R. N. (2007). Correlation based networks of equity returns sampled at different time horizons. *The European Physical Journal B*, 55(2):209–217.
- Tumminello, M., Lillo, F., and Mantegna, R. N. (2010). Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior & Organization*, 75(1):40–58.
- Vassalou, M. and Xing, Y. (2004). Default risk in equity returns. *The Journal of Finance*, 59(2):831–868.
- Zou, H. (2006). The adaptive Lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476):1418–1429.

## Appendix

In this section, we provide list of all stocks considered along with ticker symbols, name and sectoral affiliation for all four periods under study: 2002-05, 2006-09, 2010-13 and 2014-17. Finally, we also provide additional network diagrams for the full sample analysis for three periods: 2002-05, 2006-09 and 2010-13. Analysis for the data from 2014-17 is in the main text.

### A Identity of the stocks

Ticker symbols of all stocks considered: BRKA, XOM, JNJ, GE, T, PG, WMT, CVX, PFE, VZ, ORCL, KO, DIS, HD, MRK, IBM, UNH, MO, MMM, MDT, SLB, BA, LLY, HON, UPS, GS, NKE, UTX, BMY, UNP, CVS, MS, ABT, AXP, LOW, CB, BLK, COP, AIG, NEE, TMO, DHR, DUK, MET, EOG, GD, CAT, SCHW, SPG, FDX, AMT, SO, TJX, SYK, OXY, D, KMB, PRU, ITW, F, RTN, ANTM, HAL, AET, CCL, NOC, BDX, CI, EMR, PSA, MMC, ECL, APC, DE, PCG, PX, CCI, BSX, TRV, GIS, ETN, EXC, NSC, SPGI, AEP, LUV, WM, PXD, STZ, EL, MCK, HUM, ALL, VLO, TGT, APD, AFL, SHW, BAX, CBS.

In the following tables (tables 4, 5, 6 and 7), we compile the list of stocks (along with ticker symbols and sectoral affiliations) in order of centrality measure as has been discussed in the paper for four different time-periods viz., 2002-05, 2006-09, 2010-13 and 2014-17.

Table 4: Rank-ordered stocks with respect to centrality (2002-05)

Ticker	Name	Sector	Ticker	Name	Sector	Ticker	Name	Sector	Ticker	Name	Sector
ITW	Illinois Tool Works Inc	Industrials	PFE	Pfizer Inc	Health Care	ANTM	Anthem Inc	Health Care	ANTM	Anthem Inc	Health Care
AXP	American Express Co	Financials	FDX	FedEx Corp	Industrials	SPG	Simon Property Group Inc	Real Estate	SPG	Simon Property Group Inc	Real Estate
ETN	Eaton Corporation PLC	Industrials	NBE	Nextera Energy Inc	Utilities	MRK	Merck & Co Inc	Health Care	MRK	Merck & Co Inc	Health Care
GS	Goldman Sachs Group Inc	Financials	T	AT&T Inc	Telecommunication Services	PXD	Pioneer Natural Resources Co	Energy	PXD	Pioneer Natural Resources Co	Energy
EMR	Emerson Electric Co	Industrials	UPS	United Parcel Service Inc	Industrials	UNH	UnitedHealth Group Inc	Health Care	UNH	UnitedHealth Group Inc	Health Care
APD	Air Products and Chemicals Inc	Materials	F	Ford Motor Co	Consumer Discretionary	VLO	Valero Energy Corp	Energy	VLO	Valero Energy Corp	Energy
UTX	United Technologies Corp	Industrials	NKE	Nike Inc	Consumer Discretionary	BSX	Boston Scientific Corp	Health Care	BSX	Boston Scientific Corp	Health Care
ECL	Ecolab Inc	Materials	D	Dominion Energy Inc	Utilities	HUM	Humana Inc	Health Care	HUM	Humana Inc	Health Care
AIG	American International Group Inc	Financials	JNJ	Johnson & Johnson	Health Care	BAX	Baxter International Inc	Health Care	BAX	Baxter International Inc	Health Care
CAT	Caterpillar Inc	Industrials	HAL	Halliburton Co	Energy	GIS	General Mills Inc	Consumer Staples	GIS	General Mills Inc	Consumer Staples
DHR	Danaher Corp	Health Care	EXC	Exelon Corp	Utilities	PSA	Public Storage	Real Estate	PSA	Public Storage	Real Estate
SCHW	Charles Schwab Corp	Financials	AEP	American Electric Power Company Inc	Utilities	CI	Cigna Corp	Health Care	CI	Cigna Corp	Health Care
DIS	Walt Disney Co	Consumer Discretionary	AET	Aetna Inc	Health Care	EL	Estee Lauder Companies Inc	Consumer Staples	EL	Estee Lauder Companies Inc	Consumer Staples
CCL	Carnival Corp	Consumer Discretionary	CVS	CVS Health Corp	Health Care	BLK	BlackRock Inc	Financials	BLK	BlackRock Inc	Financials
MMM	3M Co	Industrials	MDT	Medtronic PLC	Health Care	AMT	American Tower Corp	Real Estate	AMT	American Tower Corp	Real Estate
HON	Honeywell International Inc	Industrials	BDX	Becton Dickinson and Co	Health Care	STZ	Constellation Brands Inc	Consumer Staples	STZ	Constellation Brands Inc	Consumer Staples
MET	MetLife Inc	Financials	RTN	Raytheon Co	Industrials	MO	Altria Group Inc	Consumer Staples	MO	Altria Group Inc	Consumer Staples
UNP	Union Pacific Corp	Industrials	KMB	Kimberly-Clark Corp	Consumer Staples	BRKk	Berkshire Hathaway Inc	Financials	BRKk	Berkshire Hathaway Inc	Financials
NSC	Norfolk Southern Corp	Industrials	WM	Waste Management Inc	Industrials	CCI	Crown Castle International Corp	Real Estate	CCI	Crown Castle International Corp	Real Estate
TRV	Travelers Companies Inc	Financials	ORCL	Oracle Corp	Information Technology						
WMT	Walmart Inc	Consumer Staples	SO	Southern Co	Utilities						
LOW	Lowe's Companies Inc	Consumer Discretionary	APC	Anadarko Petroleum Corp	Energy						
HD	Home Depot Inc	Consumer Discretionary	PG	Procter & Gamble Co	Consumer Staples						
SHW	Sherwin-Williams Co	Materials	PCG	PG&E Corp	Utilities						
TMO	Thermo Fisher Scientific Inc	Health Care	KO	Coca-Cola Co	Consumer Staples						
TGT	Target Corp	Consumer Discretionary	DUK	Duke Energy Corp	Utilities						
DE	Deere & Co	Industrials	ABT	Abbott Laboratories	Health Care						
ALL	Allstate Corp	Financials	MCK	McKesson Corp	Health Care						
SLB	Schlumberger NV	Energy	EOG	EOG Resources Inc	Energy						
LUV	Southwest Airlines Co	Industrials	GD	General Dynamics Corp	Industrials						

Table 5: Rank-ordered stocks with respect to centrality (2006-09)

Ticker	Name	Sector	Ticker	Name	Sector	Ticker	Name	Sector	Ticker	Name	Sector
UTX	United Technologies Corp	Industrials	PG	Procter & Gamble Co	Consumer Staples	RTN	Raytheon Co	Industrials			Industrials
EMR	Emerson Electric Co	Industrials	GS	Goldman Sachs Group Inc	Financials	PCG	PG&E Corp	Utilities			Utilities
DIS	Walt Disney Co	Consumer Discretionary	BLK	BlackRock Inc	Financials	CI	Cigna Corp	Health Care			Health Care
HON	Honeywell International Inc	Industrials	CCL	Carnival Corp	Consumer Discretionary	WMT	Walmart Inc	Consumer Staples			Consumer Staples
APD	Air Products and Chemicals Inc	Materials	D	Dominion Energy Inc	Utilities	MRK	Merck & Co Inc	Health Care			Health Care
PX	Praxair Inc	Materials	VZ	Verizon Communications Inc	Telecommunication Services	BMJ	Bristol-Myers Squibb Co	Health Care			Health Care
ITW	Illinois Tool Works Inc	Industrials	GE	General Electric Co	Industrials	EL	Estee Lauder Companies Inc	Consumer Staples			Consumer Staples
XOM	Exxon Mobil Corp	Energy	EXC	Exelon Corp	Utilities	MCK	McKesson Corp	Health Care			Health Care
AXP	American Express Co	Financials	MET	MetLife Inc	Financials	SHW	Sherwin-Williams Co	Materials			Materials
DHR	Danaher Corp	Health Care	CBS	CBS Corp	Consumer Discretionary	BSX	Boston Scientific Corp	Health Care			Health Care
UPS	United Parcel Service Inc	Industrials	HAL	Halliburton Co	Energy	STZ	Constellation Brands Inc	Consumer Staples			Consumer Staples
COP	ConocoPhillips	Energy	AMT	American Tower Corp	Real Estate	UNH	UnitedHealth Group Inc	Health Care			Health Care
ECL	Ecolab Inc	Materials	DE	Deere & Co	Industrials	MO	Altria Group Inc	Consumer Staples			Consumer Staples
MMM	3M Co	Industrials	ORCL	Oracle Corp	Information Technology	AET	Aetna Inc	Health Care			Health Care
LLY	Eli Lilly and Co	Health Care	NSC	Norfolk Southern Corp	Industrials	LUV	Southwest Airlines Co	Industrials			Industrials
SPGI	S&P Global Inc	Financials	NKE	Nike Inc	Consumer Discretionary	MDT	Medtronic PLC	Health Care			Health Care
ETN	Eaton Corporation PLC	Industrials	NEE	Nextera Energy Inc	Utilities	ABT	Abbott Laboratories	Health Care			Health Care
MS	Morgan Stanley	Financials	VLO	Valero Energy Corp	Energy	BAX	Baxter International Inc	Health Care			Health Care
CAT	Caterpillar Inc	Industrials	TMO	Thermo Fisher Scientific Inc	Health Care	ANTM	Anthem Inc	Health Care			Health Care
T	AT&T Inc	Telecommunication Services	UNP	Union Pacific Corp	Industrials	BDX	Becton Dickinson and Co	Health Care			Health Care
SCHW	Charles Schwab Corp	Financials	KMB	Kimberly-Clark Corp	Consumer Staples	GIS	General Mills Inc	Consumer Staples			Consumer Staples
HD	Home Depot Inc	Consumer Discretionary	WM	Waste Management Inc	Industrials	BRKa	Berkshire Hathaway Inc	Financials			Financials
PSA	Public Storage	Real Estate	GD	General Dynamics Corp	Industrials	CVS	CVS Health Corp	Health Care			Health Care
ALL	Allstate Corp	Financials	NOC	Northrop Grumman Corp	Industrials	F	Ford Motor Co	Consumer Discretionary			Consumer Discretionary
SPG	Simon Property Group Inc	Real Estate	DUK	Duke Energy Corp	Utilities	HUM	Humana Inc	Health Care			Health Care
JNJ	Johnson & Johnson	Health Care	AFL	Aflac Inc	Financials	AIG	American International Group Inc	Financials			Financials
TRV	Travelers Companies Inc	Financials	KO	Coca-Cola Co	Consumer Staples						
FDX	FedEx Corp	Industrials	SO	Southern Co	Utilities						
IBM	International Business Machines Corp	Information Technology	SYK	Stryker Corp	Health Care						
PRU	Prudential Financial Inc	Financials	MMC	Marsh & McLennan Companies Inc	Financials						

Research and Publications

Table 6: Rank-ordered stocks with respect to centrality (2010-13)

Ticker	Name	Sector	Ticker	Name	Sector	Ticker	Name	Sector	Ticker	Name	Sector
HON	Honeywell International Inc	Industrials	SYK	Stryker Corp	Health Care	MO	Altria Group Inc	Consumer Staples			
UTX	United Technologies Corp	Industrials	DE	Deere & Co	Industrials	BAX	Baxter International Inc	Health Care			
PRU	Prudential Financial Inc	Financials	TRV	Travelers Companies Inc	Financials	EL	Estee Lauder Companies Inc	Consumer Staples			
DHR	Danaher Corp	Health Care	RTN	Raytheon Co	Industrials	NKE	Nike Inc	Consumer Staples			
MMM	3M Co	Industrials	WM	Waste Management Inc	Industrials	UNH	UnitedHealth Group Inc	Health Care			
XOM	Exxon Mobil Corp	Energy	BDX	Becton Dickinson and Co	Health Care	LUV	Southwest Airlines Co	Industrials			
CVX	Chevron Corp	Energy	MS	Morgan Stanley	Financials	AMT	American Tower Corp	Real Estate			
ITW	Illinois Tool Works Inc	Industrials	SCHW	Charles Schwab Corp	Financials	SPGI	S&P Global Inc	Financials			
AFL	Aflac Inc	Financials	MDT	Medtronic PLC	Health Care	CCI	Crown Castle International Corp	Real Estate			
ETN	Eaton Corporation PLC	Industrials	JNJ	Johnson & Johnson	Health Care	ANTM	Anthem Inc	Health Care			
GE	General Electric Co	Industrials	ECL	Ecolab Inc	Materials	PG	Procter & Gamble Co	Consumer Staples			
MET	MetLife Inc	Financials	PXD	Pioneer Natural Resources Co	Energy	LLY	Eli Lilly and Co	Health Care			
EMR	Emerson Electric Co	Industrials	CCL	Carnival Corp	Consumer Discretionary	TJX	TJX Companies Inc	Consumer Discretionary			
CAT	Caterpillar Inc	Industrials	HD	Home Depot Inc	Consumer Discretionary	VZ	Verizon Communications Inc	Telecommunication Services			
MMC	Marsh & McLennan Companies Inc	Financials	D	Dominion Energy Inc	Utilities	KMB	Kimberly-Clark Corp	Consumer Staples			
PX	Praxair Inc	Materials	F	Ford Motor Co	Consumer Discretionary	SHW	Sherwin-Williams Co	Materials			
BLK	BlackRock Inc	Financials	HAL	Halliburton Co	Energy	SO	Southern Co	Utilities			
COP	ConocoPhillips	Energy	PFE	Pfizer Inc	Health Care	TGT	Target Corp	Consumer Discretionary			
GD	General Dynamics Corp	Industrials	CI	Cigna Corp	Health Care	PCG	PG&E Corp	Utilities			
UNP	Union Pacific Corp	Industrials	ORCL	Oracle Corp	Information Technology	BSX	Boston Scientific Corp	Health Care			
NOC	Northrop Grumman Corp	Industrials	KO	Coca-Cola Co	Consumer Staples	BMJ	Bristol-Myers Squibb Co	Health Care			
ALL	Allstate Corp	Financials	AEP	American Electric Power Company Inc	Utilities	WMT	Walmart Inc	Consumer Staples			
UPS	United Parcel Service Inc	Industrials	GS	Goldman Sachs Group Inc	Financials	EXC	Exelon Corp	Utilities			
DIS	Walt Disney Co	Consumer Discretionary	IBM	International Business Machines Corp	Information Technology	HUM	Humana Inc	Health Care			
BRKa	Berkshire Hathaway Inc	Financials	CVS	CVS Health Corp	Health Care	GIS	General Mills Inc	Consumer Staples			
CBS	CBS Corp	Consumer Discretionary	NEE	Nextera Energy Inc	Utilities	STZ	Constellation Brands Inc	Consumer Staples			
AXP	American Express Co	Financials	MRK	Merck & Co Inc	Health Care						
SPG	Simon Property Group Inc	Real Estate	T	AT&T Inc	Telecommunication Services						
OXY	Occidental Petroleum Corp	Energy	VLO	Valero Energy Corp	Energy						
APD	Air Products and Chemicals Inc	Materials	AIG	American International Group Inc	Financials						
CB	Chubb Ltd	Financials	DUK	Duke Energy Corp	Utilities						
PSA	Public Storage	Real Estate	MCK	McKesson Corp	Health Care						
FDX	FedEx Corp	Industrials	ABT	Abbott Laboratories	Health Care						
NSC	Norfolk Southern Corp	Industrials	AET	Aetna Inc	Health Care						
BA	Boeing Co	Industrials	LOW	Lowe's Companies Inc	Consumer Discretionary						

Table 7: Rank-ordered stocks with respect to centrality (2014-17)

Ticker	Name	Sector	Ticker	Name	Sector	Ticker	Name	Sector	Ticker	Name	Sector
BLK	BlackRock Inc	Financials	BA	Boeing Co	Industrials	T	AT&T Inc	Telecommunications Services			
ITW	Illinois Tool Works Inc	Industrials	ORCL	Oracle Corp	Information Technology	CCI	Crown Castle International Corp	Real Estate			
HON	Honeywell International Inc	Industrials	F	Ford Motor Co	Consumer Discretionary	LUV	Southwest Airlines Co	Industrials			
BRKa	Berkshire Hathaway Inc	Financials	CAT	Caterpillar Inc	Industrials	STZ	Constellation Brands Inc	Consumer Staples			
MMC	Marsh & McLennan Companies Inc	Financials	UNH	UnitedHealth Group Inc	Health Care	MO	Altria Group Inc	Consumer Staples			
AFL	Aflac Inc	Financials	MDT	Medtronic PLC	Health Care	LILY	Eli Lilly and Co	Health Care			
PRU	Prudential Financial Inc	Financials	DIS	Walt Disney Co	Consumer Discretionary	AEP	American Electric Power Company Inc	Utilities			
CB	Chubb Ltd	Financials	PFE	Pfizer Inc	Health Care	NEE	Nextera Energy Inc	Utilities			
DHR	Danaher Corp	Health Care	NSC	Norfolk Southern Corp	Industrials	VLO	Valero Energy Corp	Energy			
ETN	Eaton Corporation PLC	Industrials	SHW	Sherwin-Williams Co	Materials	D	Dominion Energy Inc	Utilities			
GS	Goldman Sachs Group Inc	Financials	WM	Waste Management Inc	Industrials	PSA	Public Storage	Real Estate			
ECL	Ecolab Inc	Materials	BDX	Becton Dickinson and Co	Health Care	EXC	Exelon Corp	Utilities			
MS	Morgan Stanley	Financials	KMB	Kimberly-Clark Corp	Consumer Staples	CI	Cigna Corp	Health Care			
MMM	3M Co	Industrials	AXP	American Express Co	Financials	WMT	Walmart Inc	Consumer Staples			
TMO	Thermo Fisher Scientific Inc	Health Care	PG	Procter & Gamble Co	Consumer Staples	MCK	McKesson Corp	Health Care			
SPGI	S&P Global Inc	Financials	MRK	Merck & Co Inc	Health Care	BMJ	Bristol-Myers Squibb Co	Health Care			
GD	General Dynamics Corp	Industrials	SLB	Schlumberger NV	Energy	DUK	Duke Energy Corp	Utilities			
UTX	United Technologies Corp	Industrials	RTN	Raytheon Co	Industrials	TGT	Target Corp	Consumer Discretionary			
APD	Air Products and Chemicals Inc	Materials	BAX	Baxter International Inc	Health Care	HUM	Humana Inc	Health Care			
EMR	Emerson Electric Co	Industrials	LOW	Lowe's Companies Inc	Consumer Discretionary	SO	Southern Co	Utilities			
PX	Praxair Inc	Materials	AMT	American Tower Corp	Real Estate						
AIG	American International Group Inc	Financials	CCL	Carnival Corp	Consumer Discretionary						
FDX	FedEx Corp	Industrials	COP	ConocoPhillips	Energy						
MET	MetLife Inc	Financials	HAL	Halliburton Co	Energy						
SCHW	Charles Schwab Corp	Financials	GHS	General Mills Inc	Consumer Staples						
TRV	Travelers Companies Inc	Financials	EL	Estee Lauder Companies Inc	Consumer Staples						
UPS	United Parcel Service Inc	Industrials	AET	Aetna Inc	Health Care						
ABT	Abbott Laboratories	Health Care	KO	Coca-Cola Co	Consumer Staples						
XOM	Exxon Mobil Corp	Energy	CBS	CBS Corp	Consumer Discretionary						
JNJ	Johnson & Johnson	Health Care	APC	Anadarko Petroleum Corp	Energy						
UNP	Union Pacific Corp	Industrials	VZ	Verizon Communications Inc	Telecommunications Services						
NOC	Northrop Grumman Corp	Industrials	BSX	Boston Scientific Corp	Health Care						
HD	Home Depot Inc	Consumer Discretionary	CVS	CVS Health Corp	Health Care						
CVX	Chevron Corp	Energy	NKE	Nike Inc	Consumer Discretionary						
SYK	Stryker Corp	Health Care	DE	Deere & Co	Industrials						

## B Additional network figures

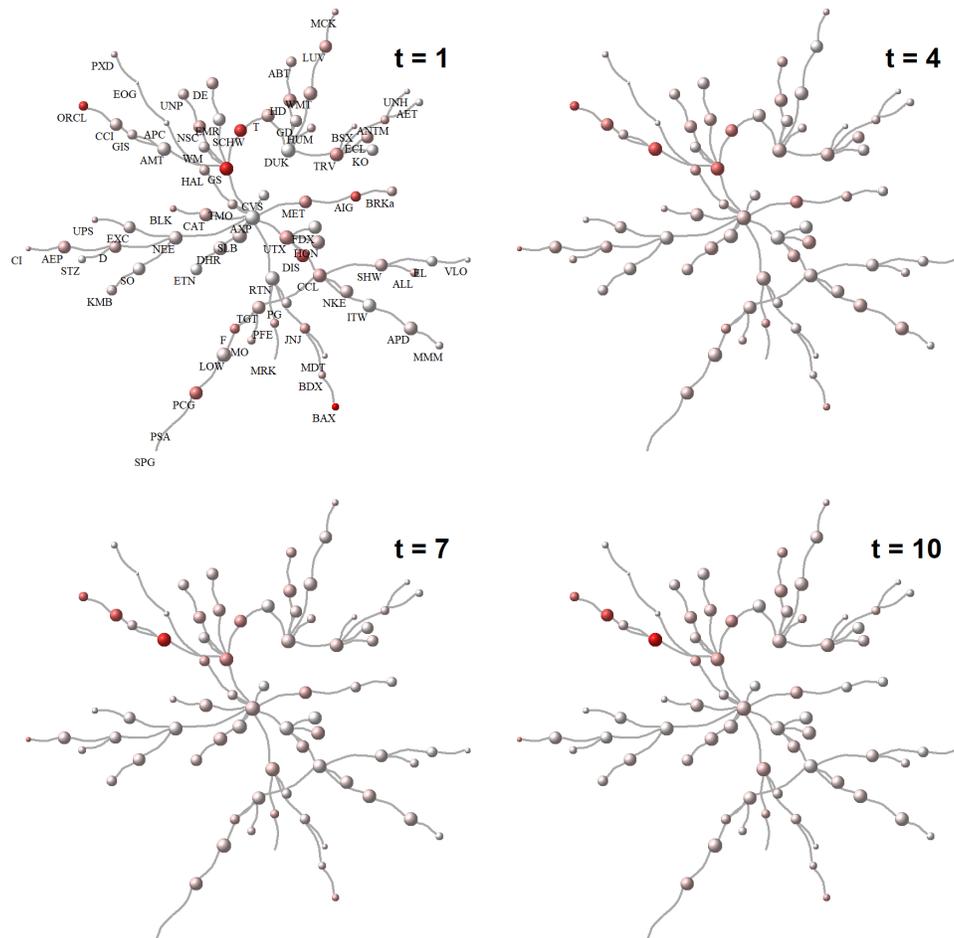


Figure 5: (color online) Ripple effects on the network of top 100 firms in terms of market capitalization during the period 2002-05. We selected the maximally connected component by Lasso estimation. The epicenter of the shock is Goldman Sachs (GS). The diagram exhibits responses of the stocks at four time points  $t = 1$  (left-top panel), 4 (right-top panel), 7 (left-bottom panel) and 10 (right-bottom panel). Intensity of color represents the magnitude of return volatility of the nodes in the network.

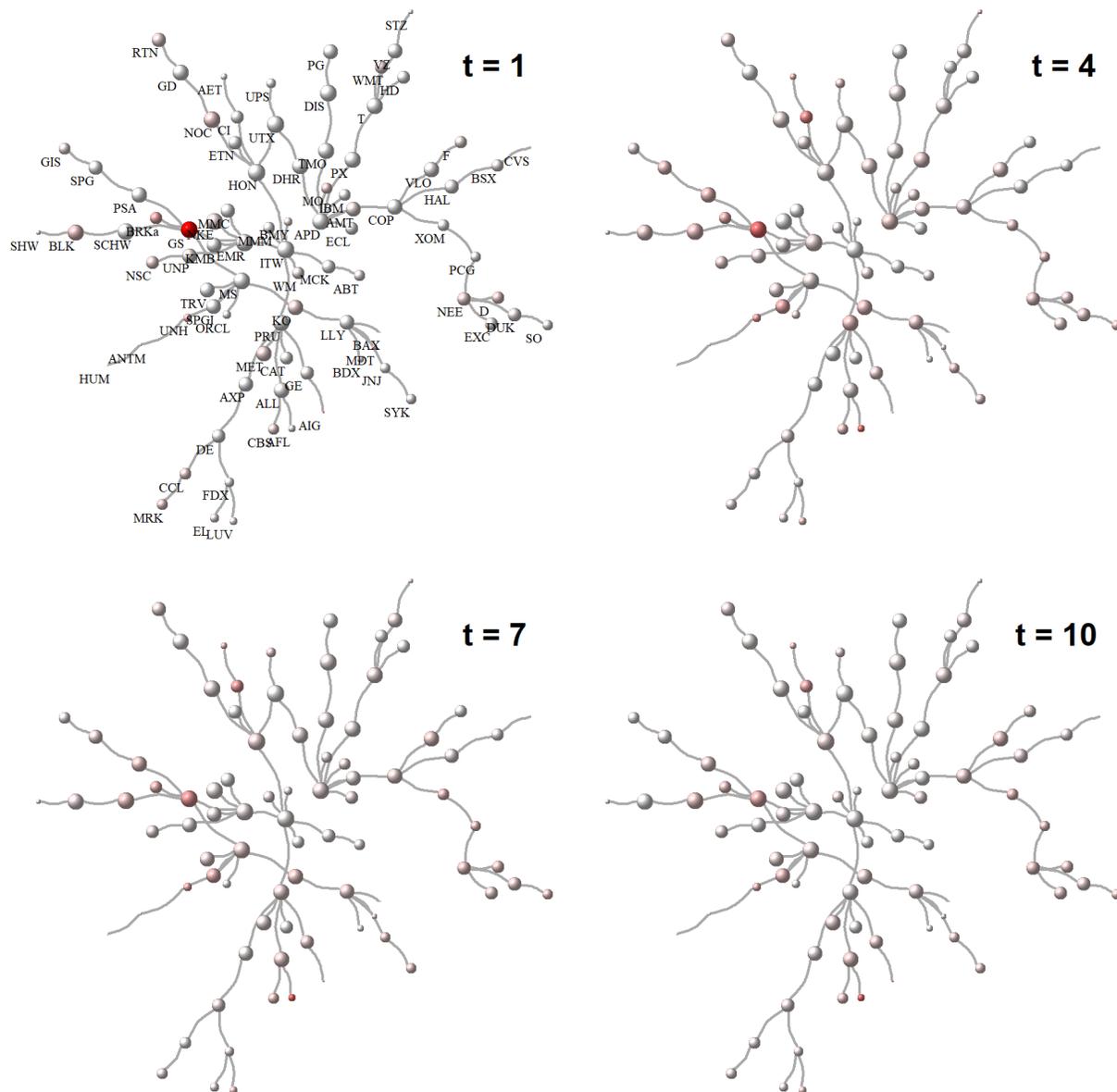


Figure 6: (color online) Ripple effects on the network of top 100 firms in terms of market capitalization during the period 2006-09. We selected the maximally connected component by Lasso estimation. The epicenter of the shock is Goldman Sachs (GS). The diagram exhibits responses of the stocks at four time points  $t = 1$  (left-top panel), 4 (right-top panel), 7 (left-bottom panel) and 10 (right-bottom panel). Intensity of color represents the magnitude of return volatility of the nodes in the network.

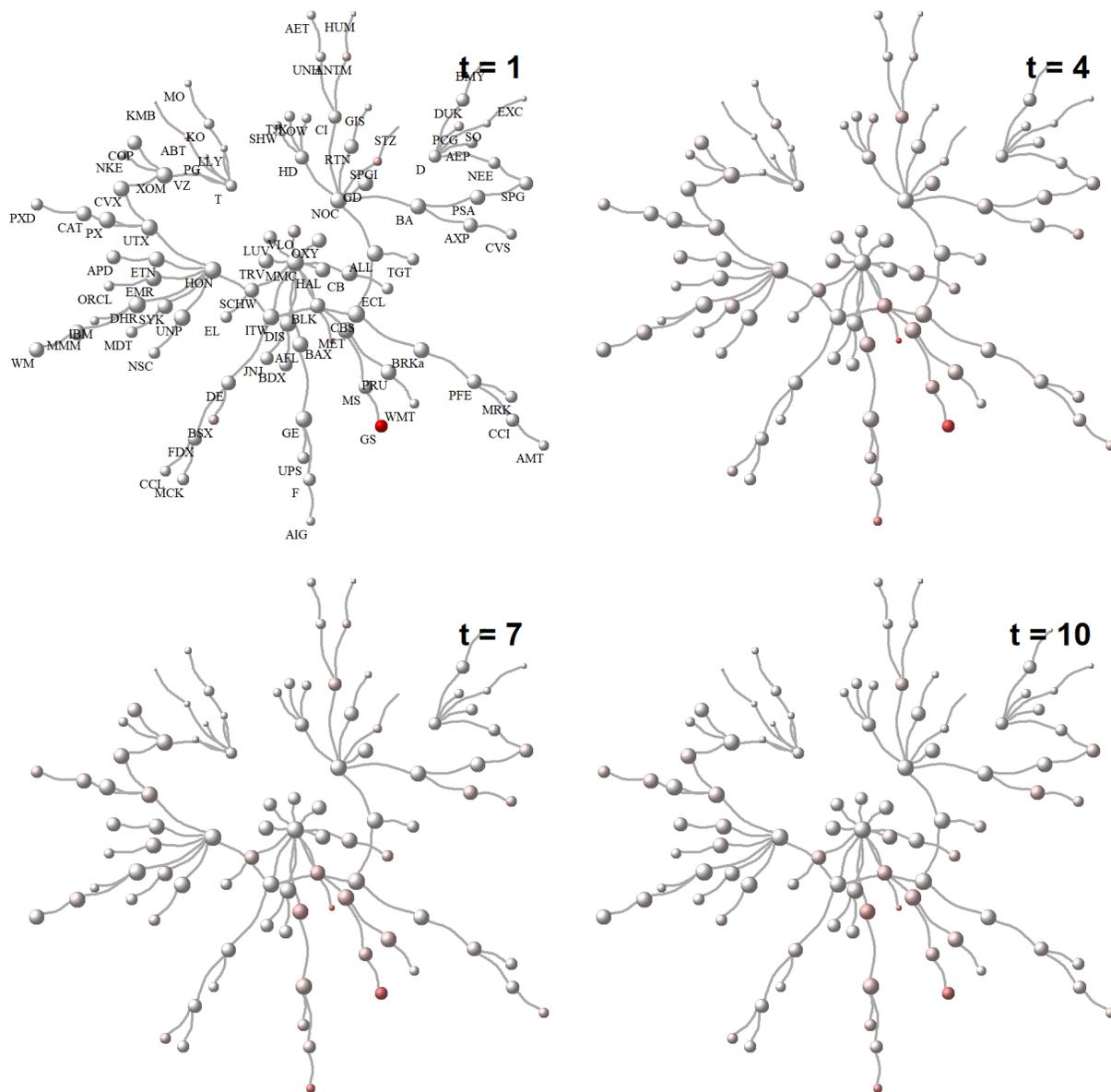


Figure 7: (color online) Ripple effects on the network of top 100 firms in terms of market capitalization during the period 2010-13. We selected the maximally connected component by Lasso estimation. The epicenter of the shock is Goldman Sachs (GS). The diagram exhibits responses of the stocks at four time points  $t = 1$  (left-top panel), 4 (right-top panel), 7 (left-bottom panel) and 10 (right-bottom panel). Intensity of color represents the magnitude of return volatility of the nodes in the network.