


Ripples on financial networks

Sudarshan Kumar , Avijit Bansal & Anindya S. Chakrabarti

To cite this article: Sudarshan Kumar , Avijit Bansal & Anindya S. Chakrabarti (2020): Ripples on financial networks, The European Journal of Finance

To link to this article: <https://doi.org/10.1080/1351847X.2020.1835686>




View supplementary material 



Published online: 20 Oct 2020.



Submit your article to this journal 



View related articles 



View Crossmark data 



Ripples on financial networks

Sudarshan Kumar^a, Avijit Bansal^b and Anindya S. Chakrabarti^c

^aFinance, Accounting and Control, Indian Institute of Management, Kozhikode, India; ^bFinance & Accounting Area, Indian Institute of Management, Ahmedabad, India; ^cEconomics Area, Indian Institute of Management, Ahmedabad, India

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 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 estimate 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.

ARTICLE HISTORY

Received 20 September 2019
Accepted 28 September 2020

KEYWORDS

Volatility spillover; VAR identification; spectral analysis; core-periphery

JEL CLASSIFICATIONS


C55; C58; G17; C53

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 asset prices do not evolve over time in silos. They are affected by and in turn affect other asset prices. Thus the quantification of the effects of a volatility shock emanating from a given epicenter percolating through the network would depend on the architecture of the network denoting linkages across assets. In this paper, we propose a unique rank-order of assets obtained from the network topology and utilize a many-dimensional vector autoregression model identified by the rank-ordering, to characterize the shock spillovers or *ripples* on large-scale financial networks.

There has been significant 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 (Tumminello et al. 2007; Tumminello, Lillo, and Mantegna 2010; Acemoglu et al. 2012; Pozzi, Di Matteo, and Aste 2013; Acemoglu, Ozdaglar, and Tahbaz-Salehi 2015; Diebold and Yilmaz 2015a; Barigozzi and Hallin 2017a, 2017b; Dessaint et al. 2018; Geraci and Gnabo 2018, among others). 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, Ozdaglar, and Tahbaz-Salehi (2016) for a review of the literature). A complementary literature on the financial markets focused on different forms of the shock propagation mechanism. 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,*

CONTACT Anindya S. Chakrabarti  anindyac@iima.ac.in

 Supplemental data for this article can be accessed here. <https://doi.org/10.1080/1351847X.2020.1835686>

their customers, and other markets'. In this paper, we focus on the nature of interconnectivity and model the phenomenon of shock propagation from an econometric point of view where one has to estimate the network from underlying process and data, in line with the reduced-form approach proposed 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 comprises own dynamics as well as impact of other assets' dynamics. Thus the identification of shock spillover cannot be deduced by considering stocks separately. We accommodate the information of joint evolution by constructing a network of stocks and for modeling the time domain information, we utilize a vector autoregression (VAR henceforth) framework with an identification restriction obtained from the underlying network topology. This leads to the second problem about 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. 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 heteroscedastic (GARCH henceforth) framework. We allow the lags to vary within a range and by using information criteria, we select the best lag combinations for each of the stock return series. In order to analyze shock propagation, we estimate a VAR model on the inferred conditional volatility series. The VAR model is identified by imposing an ordering on the stocks obtained the eigenvector centrality of the observed return correlation matrix (denoted by c_r^{eig} henceforth). The essential idea of the identification is that volatility shocks would percolate from high centrality stock stocks to low centrality stocks.

This identification choice reflects two mechanisms. First, the dominant eigenvalue of the observed return correlation matrix represents the *market factor* (Plerou et al. 2002), i.e. contribution of each asset to the aggregate market volatility. Thus a high value of a stock in the dominant eigenvector represents higher participation of the stock, which provides information about the relative exogeneity of assets in the market. Our identification restrictions rely on the mechanism that volatility spillovers occur from a more exogenous stock to relatively less exogenous stocks. The dominant eigenvector is a natural choice for such ordering criteria. Second, we observe in this connection that the dominant eigenvector has an alternative interpretation in network theory. It represents eigenvector centrality (EVC henceforth) which produces a recursive measure of importance of the nodes in a graph. As we will explain later in the paper, eigenvector centrality is also the limiting case of Katz–Bonacich centrality which is a generalized class of centrality measures, 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. 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 epicenter on the network.

Since the eigenvector centrality of the return correlation matrix (c_r^{eig}) is the fundamental object in the proposed methodology, we provide analytical as well as empirical justifications for using c_r^{eig} as the identification criteria. In principle, one could also use the eigenvector centrality of the volatility correlation matrix (c_v^{eig} henceforth) as an ordering criteria for identification. However, we provide two strong empirical rationales 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 4 equally non-overlapping wide windows ranging over 4 years viz. 2002–05, 2006–09, 2010–13 and 2014–17. These windows correspond to pre-crisis, crisis, post-crisis and relatively calm periods respectively. We observe the evolution of c_r^{eig} in these four time window (Table A1). It is evident that the correlation values of pairs of c_r^{eig} across time periods are greater than 0.5 in five out of six cases. In comparison, the corresponding values for c_v^{eig} are considerably lower (Table A2). The second rationale is that the definition and measurement of return are unique and parameter-independent 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. Hence, identification criteria based on the ordering of volatility centrality would be sensitive to model specification and therefore, non-robust.

In comparison, c_r^{eig} is uniquely defined and hence, more credible as an identification criteria. In this context, it may be noted that c_r^{eig} is a purely statistical measure based on asset return comovements. We use the bounds of *Marčenko–Pastur* distribution from random matrix theory (Marenko and Pastur 1967), to show that c_r^{eig} corresponds to a statistically significant eigenvalue.

For the purpose of visualization, we exploit *minimum spanning tree* extracted from the full scale network. In the methods section (see Section 3), we discuss the benefits of this choice. We estimate a many-dimensional vector autoregressive model on the conditional volatility series with identification restrictions implied by the return network. Exploiting the resultant VAR structure, we generate the impulse response functions which we plot on the minimum spanning tree to demonstrate the ripples of shock propagation across stocks. In Section 6, we discuss a way to increase efficiency of the process by reducing dimensionality 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.

Finally, we explore the origin of variations of stocks in c_r^{eig} , which constitutes the identifying restrictions. In particular, 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 the number of analysts (see Section 5.2). The coefficient signs of the control variable in the panel regression also have reasonable interpretation. More profitable firms tend to exhibit higher centrality and firms with bad ratings tend to have lower centrality. Earlier work had also noted that for sectoral indices, c_r^{eig} is strongly correlated with the size of the sector measured by market capitalization, revenue as well as employment (Sharma et al. 2017). However, the present result establishes the result on a much larger dataset and the results are consistent across time.

The rest of the paper is structured as follows. In the next section, we review the relevant literature. In Section 3, 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 Sections 5 and 6, we provide justification of the identifying restrictions and a mechanism to increase efficiency of the algorithm via Lasso estimator, respectively. Section 7 summarizes the paper and concludes.

2. Background literature

Our paper is part of a growing literature that analyzes the mechanism of shock propagation on financial networks. However, 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 (Mantegna 1999; Tumminello, Lillo, and Mantegna 2010; Pozzi, Di Matteo, and Aste 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) contain 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 doubtful. 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 questionable. 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 heteroscedasticity* or GARCH model to estimate the volatility process. Other modeling techniques like the one proposed by Engle and Figlewski (2014) uses an 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 (which in any case cannot be applied for a large number of time-series). This approach is quite prevalent in the literature on shock propagation (Diebold and Yilmaz 2015a). Apart from identification, we differ significantly from the above literature in terms of network construction. We utilize return correlation-based network rather than using network constructed from long-range variance decomposition (Mantegna 1999).

The idea that there are spillover 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.

It is important to summarize the differences of our work from the line of work initiated by Diebold and Yilmaz (2015a); Barigozzi and Hallin (2017a). The main methodological difference boils down to the fact that this literature typically relies on variance network computed from forecast errors. In comparison, our proposal is to consider the network generated only from the realized time series of asset prices. A more important analytical point of departure is the usage of network topology in this paper to uniquely identify the paths of shock spillovers. Taken together, we provide a complete method to compute *ripples* on the financial networks.

Finally, we observe that the literature on graph theoretic treatment of VAR models is very limited. Here we discuss the differences between partial correlation based identification from our approach. Killian and Lütkepohl described (see the discussion in p. 239, Kilian and Lütkepohl (2017)) a graph theoretic approach that relies on finding pairwise partial correlation and exclusion restriction is to impose zero for covariance if the partial correlation is statistically insignificant. Our approach differs in that we are proposing a recursive identification rather than a non-recursive scheme that utilizes only binary relationships (partial correlation) rather than the full structure of the network. Therefore, although this approach is described as a graph theoretic approach, our proposed mechanism actually exploits the graph/network structure. In the present context, Kumar, Di Matteo, and Chakrabarti (2020) is a recent attempt that proposed an identification criteria based on the topology of network. However, they differ explicitly in terms of proposing a non-recursive identification scheme based on the properties of graph planarity. Our approach does not rely on the planar graph structure and more generally applicable with a relatively straight-forward recursive identification scheme.

3. Empirical data and methodology

In the following, we provide the complete description to characterize the shock propagation across the asset network. In subsequent discussions, background and usage of each of the methods have been described in details. In Section 3.4, we provide a concise set of steps that encapsulates the entire analysis.

3.1. 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).¹ We denote the entire time period by T . All stocks are chosen such that

they have existed throughout the entire time period. We divide the 16 years long period into 4 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). 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 window captures the time period after crash when the economy was still recovering and the fourth window captures a period of relatively higher stability.

3.2. Return and volatility series

Let us denote the set of stocks by \mathcal{N} which has cardinality of N . Daily closing stock prices are denoted by $\{p_{it}\}_{i \in \mathcal{N}, t \in T_j}$ 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 $i \in \mathcal{N}$, $t \in T_j$ 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.²

Estimation of conditional volatility: We use a simple 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 for the i th asset, that

$$\begin{aligned} r_{it} &= \mu_i + \sigma_{it}\epsilon_{it} \\ \sigma_{it}^2 &= c_i + \sum_{l=1}^p \alpha_{il} r_{i,t-l}^2 + \sum_{j=1}^q \beta_{ij} \sigma_{i,t-j}^2. \end{aligned} \quad (2)$$

We will denote conditional volatility of the i th stock by σ_{it} .³

Cross-correlation matrix: We construct a sample correlation matrix Γ^x from both the return series $\{r_{it}\}$ and the volatility series $\{\sigma_{it}\}$, by using pairwise correlation coefficients defined as

$$\gamma_{ij}^x = \frac{E((x_{it} - \bar{x}_i)(x_{jt} - \bar{x}_j))}{\sigma_i^x \cdot \sigma_j^x} \quad (3)$$

for $i, j \in n$, where \bar{x}_i and σ_i^x respectively denote mean and standard deviation of the variable x_{it} which can represent return (r_{it}) or conditional volatility (σ_{it}). We denote the equal time cross-correlation matrices by $\Gamma^r = \{\gamma_{ij}^r\}$ and $\Gamma^\sigma = \{\gamma_{ij}^\sigma\}$ respectively.

3.2.1. Identification restrictions from network topology

Once, we have the estimated latent volatility, 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 based on the cross-correlation matrix Γ described above. We consider an unique criteria, viz. the limiting Katz-Bonacich centrality (Bonacich 1987) which is defined as follows.

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

$$c^{KB}(\alpha) = \left((I - \alpha A^T)^{-1} - I \right) \mathbf{I}, \quad (4)$$

where V is a set of nodes, E denotes a set of edges, α is an attenuation factor (imposed exogenously), I is an identity matrix and \mathbf{I} is a vector of ones. Under the condition that $\alpha < 1/\lambda_{\max}$ (λ_{\max} being the highest eigenvalue of A), the centrality measure is well-defined.

For uniqueness of the measure and simplification of exposition of the empirical analysis, we consider the ranking produced by limit of the ranking produced by Katz-Bonacich centrality which converges to eigenvector centrality⁴ c^{eig} , as stated in the result below.

Result [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^{ig} .

The relative exogeneity of the stocks can be defined as having highest centrality in the correlation matrices Γ^r and Γ^σ . In this paper, we will empirically demonstrate that centrality computed from Γ^r is a substantially better exogeneity criteria due to its stability over all time periods (see Section 5). 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 based on eigenvector centrality of the stocks obtained from the correlation matrix.⁵

3.2.2. Vector autoregression estimation and orthogonalization

Now, we estimate a vector autoregression model with p -lags of the following form⁶:

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + u_t \quad (5)$$

where all matrices $\{A_i\}_{i=1,\dots,p}$ have dimensions of $N \times N$, on the estimated volatility series $\{\sigma_{it}\}_{i \in \mathcal{N}}$. The noise term has a covariance matrix $E(u_t u_t^T) = \Xi$.

We find that all information criteria chooses a lag of 1 (except AIC which chooses 2, consistent with the nature of the lower penalty associated with the lags; we have used *vars* package in *R*) in the empirical analysis. 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). We also note that more lags would not necessarily provide any economic or financial 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 non-trivial model capturing the comovement 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 ξ , i.e.

$$\hat{\Xi} = \xi \xi^* \quad (6)$$

where ξ^* denotes the conjugate transpose of ξ . Then following standard protocol, we find the orthogonalized impulse response functions by pre-multiplying Equation (5) with the matrix ξ .

3.3. 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 Figure 1 (discussion on this figure can be found in Section 4.1). To enhance visual clarity, we have utilized the minimum spanning tree (see Section 3.3.1) and depicted the propagation of volatility shock emanating from a chosen epicenter. We discuss visualization in more details below.

3.3.1. Visualization

If we plot a network with N stocks, the number of possible edges would be N^2 . For large values of N , such a large number of edges will make it impossible to visualize the network and make any sense of it. Therefore, we need a clean and unique way to visualize that will retain a subset of connections between the stocks and still be useful for visual exposition. To achieve such an objective, we propose constructing a *minimum spanning tree* (MST hereafter) on the equal-time cross-correlation matrix. We will show that it provides a clean and unique presentation of the network such that the network remains connected with smallest number of edges.⁷ Below we describe the methodology in details.

Recall that in Section 3.2, we have constructed sample correlation matrices Γ^x from both the return series ($x = \{r_{it}\}$) and the volatility series ($x = \{\sigma_{it}\}$) in Equation (3). We note that individual correlation values cannot serve the purpose of distance as they can be negative. Therefore, following the literature (Mantegna 1999;

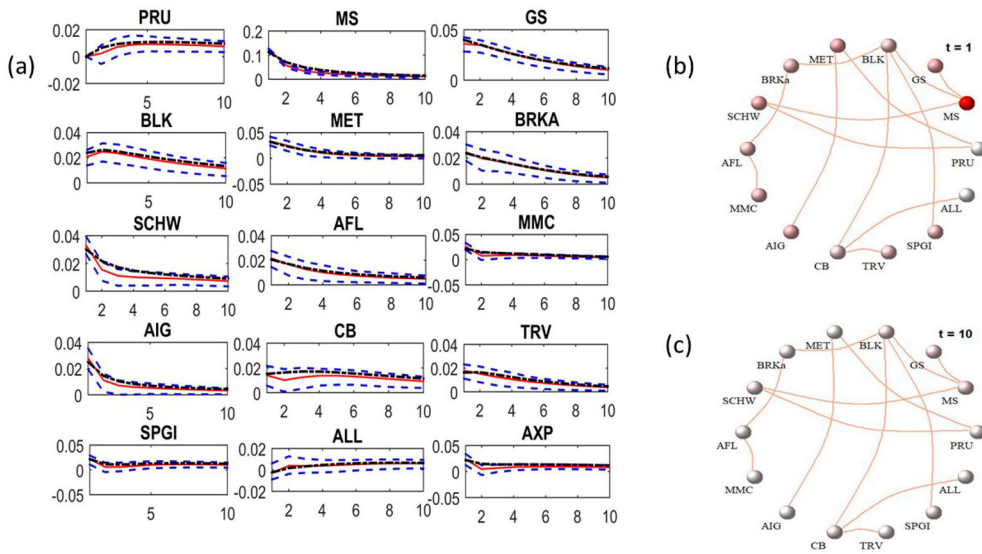


Figure 1. Example of ripple effects on the financial network during the period 2014–17. We have chosen all financial firms within the set of 100 firms with largest market capitalization in the NYSE, to construct the network. *Panel (a):* Impulse response functions along with 90% confidence intervals across different stocks, in reaction to a positive volatility shock given to Morgan Stanley (MS). The black dashed line shows average impulse response function with parameters bootstrapped from the GARCH estimation over 10^3 times (see text for more details on the calculations). *Panel (b and c):* Shock propagation shown across the network among the stock (for visualization we have utilized a minimum spanning tree) at two time points $t = 1$ (*panel (b)*) and 10 (*panel (c)*). Intensity of color represents the magnitude of return volatility of the nodes in the network.

Bonanno et al. 2004) we consider the metric due to Gower (1966):

$$d_{ij} = \sqrt{2(1 - \gamma_{ij})}. \quad (7)$$

Using this nonlinear transformation, we can construct a distance matrix D^r from the correlation matrix Γ^r which gives rise to a network $G^r = (V^r, E^r)$ where V denotes the set of nodes/vertices (stocks) and E denotes the set of edges. Clearly this network contains 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* from Γ^r . We will utilize the MST to graphical exposition of the ripple effect as it drastically reduces the number of links or edges to $n-1$. Formally the definition is as follows: 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.⁸

3.4. Algorithm to characterize the ripple effects

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

- (1) We choose N number of stocks to create return series over T time periods⁹ such that $T \gg N$.
- (2) Next, we estimate latent volatility series from each of the return series by using *generalized autoregressive conditional heteroscedasticity* model.
- (3) We estimate a vector autoregression model on the log of latent volatility series across all stocks in the sample. The VAR model is identified through Cholesky decomposition of the error covariance matrix, by imposing an ordering on the stocks obtained from the eigenvector centrality c_r^{eig} of the return correlation matrix.
- (4) Now one can characterize the shock propagation over the stocks by using estimated impulse response functions obtained the identified VAR model.

- (5) For visualization of shock propagation, we construct a network (G') based on the correlation structure of the stocks and extract the minimum spanning tree.¹⁰ We plot the impulse responses across the network emanating from chosen epicenters.

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

4. Empirical analysis

We implement our methodology on two sets of data for illustrative purpose. First, we analyze spill over effects over a small set of financial firms. Second, we apply the technique 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.

4.1. Example: spill over effects on the financial firms

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 stocks during the period 2014–17. In particular, we analyze the results of a positive volatility shock given to Morgan Stanley (MS).¹¹

Figure 1 shows the responses of an unit orthogonalized volatility shock emanating from Morgan Stanley (MS) to all other financial stocks (which belong to the set of top 100 stocks in New York stock exchange in terms of market capitalization). Panels on the left (under *panel (a)*) show complete descriptions of the impulse response functions across periods t ranging from 1 to 10 for 15 stocks (PRU in *top-left* to AXP *bottom-right*), along with 90% confidence intervals (blue dashed lines). This representation depicts the standard description of impulse response functions. However, the VAR estimates can be susceptible to the distribution of the GARCH parameter estimates since latent volatility process is a *point estimation*. To take care of that, we simulate impulse response functions obtained from bootstrapped GARCH parameter estimates of the latent volatility time series. In particular, we simulate 10^3 number of latent volatility series based on each estimated GARCH parameter distribution, for all 100 stocks. Then we estimate vector autoregression model over 10^4 independent sample draws with replacement from the population of simulated latent volatility series. We see that the simulated impulse response functions averaged over these 10^4 estimations fully lie within the 90% confidence interval and in most of the cases, the difference from the original impulse response function is minimal.¹²

In *panels (b)* and *(c)*, we convert these figures into a network-oriented depiction. We show the network of stocks at $t = 1$ (time when the volatility shock hits; *panel (b)*) and the same at $t = 10$ (*panel (c)*). The color intensity denotes the strength of the shock spill over across nodes of the network. We see that Charles Schwab Corp. (SCHW), Berkshire Hathaway Inc. (BRKA), MetLife Inc. (MET), BlackRock Inc. (BLK) and Goldman Sachs Group Inc. (GS) were immediately affected. But after 10 periods, most of them have recovered with some residual impacts on BLK and GS.

In the following, we only consider the network depiction. For each result, we can provide the standard representation of the impulse response functions. But the network depiction allows us to visualize the shock propagation in a clean manner.

4.2. 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.

In the first experiment, 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 Morgan Stanley (MS) 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. In the second experiment, 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.

Figure 2 shows the ripples emanating from Morgan Stanley (MS) in response to a positive volatility shock in the period 2006–09. We have followed the entire procedure described in Section 3.4. The diagram shows four snapshots at time $t = 1, 4, 7$ and 10. We see that firms like Goldman Sachs (GS), MetLife Inc. (MET) and Charles Scwab Corp (SCHW) among others are affected. The statistical results are consistent with the economic intuition that mostly financial stocks respond to the shock to MS which is financial firm. Responses to unit shock emanating from the same epicenter Morgan Stanley (MS) in different time windows have been shown in

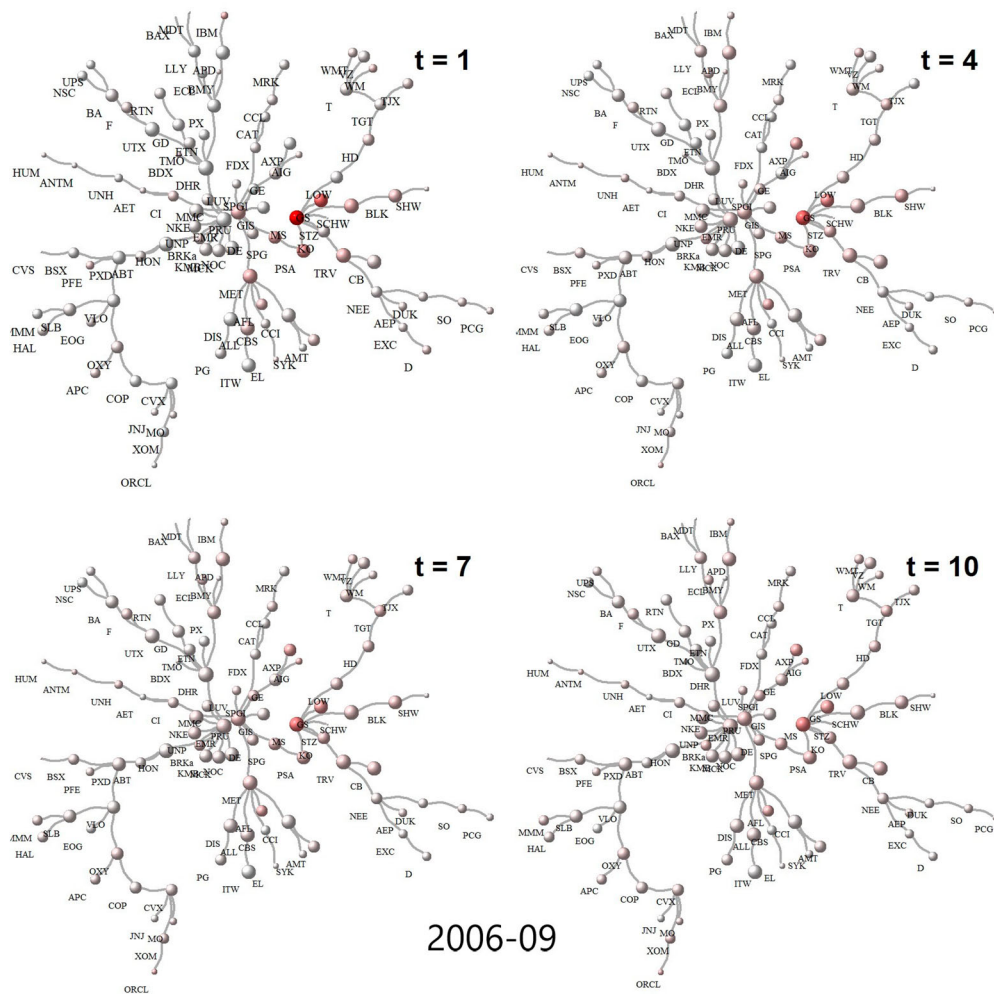


Figure 2. Ripple effects over 10 days ($t = 1$ (left-top panel); 4 (right-top panel); 7 (left-bottom panel); 10 (right-bottom panel)) on the network of top 100 firms in terms of market capitalization during the period 2006–09. The epicenter of the shock is Morgan Stanley (MS). Intensity of color represents the magnitude of return volatility of the nodes in the network.

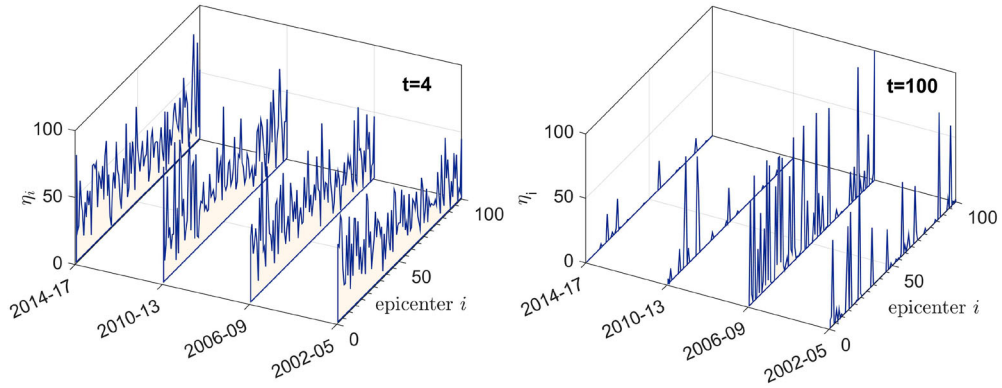


Figure 3. A summary of impulse response functions (z-axis) across all stocks (y-axis) across all four periods (x-axis). On the z-axis we plot the number of stocks showing at least 5% response (in absolute values) with respect to a shock given to the epicenter i (listed on the y-axis, where $i = 1$ to 100) at $t = 4$ (left panel) and at $t = 100$ (right panel). The spillover intensity of the shock clearly decline over time. But the persistence in the period 2006–09 is noteworthy.

Figures A1, A2, A3, in the Appendix. Interestingly, as the figures suggest the effects were much stronger during the period 2006–09, and the effects were substantially subdued during the rest of the periods.

Although the first experiment shades light on the phenomenon on shock spillover, it does not provide an aggregate picture since 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 conduct the second experiment where we summarize the information of the impact of individual stocks on the remaining market in Figure 3. 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 (arranged in terms of centralities) which act as the epicenters of the ripple effects.

In successive rounds, we give unit shocks to all the stocks treating them as epicenters and trace the ripples emanating from them spreading throughout the network. In order to see the impact of shocks, we plot on the z-axis the number of firms showing response of at least 5% magnitude of the shock given to the epicenter, after 4 time points i.e. $t = 4$ (see Figure 3, left panel). Across four snapshots, we see very similar degree of average responses (32, 34, 35, 33 respectively).

The right panel in Figure 3 shows the same response computed after 100 periods ($t = 100$). Two features can be noticed. First, there is a sharp decline in the number of stocks showing responses for all epicenters. Second, a large number of stocks in the period 2006–09 still shows comparatively high response indicating that during the time of financial crisis, the effects of the shock remained more persistent than during the other periods.

5. Return centrality as identification criteria

We note that in Section 3, we introduced the criteria for identifying the VAR model with a purely data-driven justification. To summarize, we have shown that the stocks can be uniquely rank-ordered with respect to eigenvector centrality (or limiting Katz–Bonacich centrality) and it provides a natural interpretation of shock propagation from a network theoretic point of view. In this section, we discuss the financial and econometric justifications for considering return centrality as the identification criteria.

First, we show that eigenvector centrality has a clear interpretation in terms of presence in a *market portfolio* as explained below. The main motivation comes from eigenvalue decomposition of cross-correlation matrix of return series. Let us denote the cross-correlation matrix by Σ_r . We can decompose the matrix as

$$\Sigma_r = \sum_{i=1}^n \lambda_i c_i c_i^T \quad (8)$$

where λ_i is the i th eigenvalue and c_i is the i th eigenvector. Let us denote the highest eigenvalue in modulus is λ_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).

In order to utilize the eigenvector centrality as a statistically credible identification criteria, we need to make sure that the dominant eigenvalue λ_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. We utilize a result from Marenko and Pastur (1967) that goes as follows: 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

$$\lambda_{u.b.} = \left(1 + \frac{1}{\sqrt{Q}}\right)^2. \quad (9)$$

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

In the following, we demonstrate two points. First, we show that eigenvector centrality computed from the return correlation matrix is more robust than that computed from the volatility correlation matrix (see Section 5.1). Second, we show that return centrality is related to firm-level fundamentals (see Section 5.2).

5.1. Centrality measures: stability over volatile and tranquil periods

Here, we analyze the stability of the rank-ordering obtained from for two types of identification criteria. In principle, 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. However, 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 A1 in four different snapshots (2002–05, 2006–09, 2010–13 and 2014–17) that covers pre-crisis to post-crisis periods. In Table A2, we show the same for volatility correlation centrality.¹³ Return centralities are clearly much more correlated across time than the corresponding volatility centralities. Thus the return centrality measure is more robust.

5.2. Explanatory factors of return centrality c_r^{eig}

So far we have utilized the return centrality c_r^{eig} as an identification criteria for finding the impulse response functions from the VAR model and in Section 5.1, we have established that the implied rank-order is more stable than that obtained from volatility centrality. In this section, we explore its relationship with firm-level fundamentals and show that economic size explains return centrality among others.

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). We are unaware of any literature that relates firm-level economic or financial 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, Liu, and Zhu (2016) provide 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 A3 gives the result of these regression for all four time windows (2002–06, 2006–09, 2010–13 and 2014–17). Log assets are generally strongly correlated with the centrality measure with positive coefficients. An immediate implication is that shock propagates from large firm to smaller firms (under our proposed ordering assumption in the VAR model). 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. 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; see Table 4 in the Online Supplementary Material) with time effect retains the feature that firms having higher centrality in the financial network, are larger in size, more profitable and are in better 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.

6. Efficiency increase through dimension reduction (Lasso)

We note that the methodology described above is generally applicable to any set of stocks. Therefore, to estimate the model there has to be commensurate amount of data, since the proposed algorithm does not penalize the system size, i.e. the number of stocks analyzed. In many cases, it would be useful to reduce the system size, for example due to data limitations or if we are interested in analyzing only the strongly connected subcomponent of the network in which the spillover would be most prominent.

In order to address such issues, we first need to identify the set of stocks which can be affected by the shock propagation (Hsu, Hung, and Chang 2008). Towards that objective, we first construct *maximally connected component* G_{mcc} of the network G . We define such a component as follows: 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:

$$\min \sum_{t=1}^T \left(y_t - \sum_{j=1}^k \beta_j x_{jt} \right)^2 \quad \text{s.t.} \quad \sum_{j=1}^k |\beta_j| \leq C \quad (10)$$

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 $\{\sigma_{it}\}_{i \in \mathcal{N}, t \in T_j}$ 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 β . Then we convert into a binary matrix (0, 1) where all nonzero β_{ij} are converted into ones and the β_{ij} which were assigned a value of zero by the Lasso estimation, remain zero. The row sum and the column sum of this binary transformation of the β matrix represents *indegree* and *outdegree* respectively.¹⁴ Then we use a threshold to differentiate between connected and unconnected nodes. As demonstrative empirical exercises, we carried out the size reduction through LASSO following the convention that if a node has both *indegree* and *outdegree* less than $f\%$ where f is chosen exogenously, 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 $f\%$ 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 $n \leq N$. In our datasets, we find that with $f = 10$, i.e. a 10% threshold, the number of surviving stocks are 80 (2002–05), 97 (2006–09), 95 (2010–13) and 94 (2014–17) respectively over 4 snapshots out of 100 stocks. If we increase the threshold to 15%, then the sequence of the number of surviving stocks is 43, 77, 78 and 68. A further increase to 20% leads to the number of surviving stocks in the connected network as 18, 39, 48 and 30. Finally, we would like to add an observation here that given the numerical nature of the optimization on a high-dimensional object, sometimes convergence is not obtained as would be expected out of a numerical routine for large-dimensional optimization (we have used *sparsevar* package in R). The exact numbers of stocks that are selected show minor variations with respect to the tuning parameters inclusive of number of folds used for cross validation, the shrinkage penalty as well as the method of optimization.

7. Summary and conclusion

Propagation of shocks on a financial network is an intuitive phenomenon that has attracted substantial attention in the recent literature on financial econometrics. The standard approach is to analyze the long range variance decomposition. In this paper, we propose a new 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 the dominant eigenvector of the observed return correlation matrix provides an intuitive and robust identification strategy. The identified model allows us to observe and uniquely characterize the ripple effects originating from a particular epicenter in the stock network and to quantify the corresponding magnitude of fluctuations.

We show that 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. In principle, the network can be decomposed into a core and periphery, but centrality provides a continuous version of relative coreness than an ad hoc binary characterization of nodes being either in core or in periphery. Second, this feature becomes very useful for defining relative exogeneity of the stocks, which acts as an identification criteria for finding orthogonalized impulse response functions through standard Cholesky decomposition. Third, we utilize a result from random matrix theory (in particular, the celebrated Marcenko–Pastur theorem developed by Marcenko and Pastur (1967)) to establish the statistical validity of the centrality measure that the result is not an outcome of random noise. This approach was used by Plerou et al. (2002) and has become useful for analyzing financial risk with

high-dimensional data (see, e.g. Bouchaud and Potters (2003)). We also empirically document that the proposed return centrality has a strong relationship with firm-level fundamentals including size.

Modeling spillover through linkages of volatilities has 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. 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); Hue, Lucotte, and Tokpavi (2019) among others, highlight the role of networks in measurement and contribution to systemic risk. Other strands of work relate the network structure of the financial markets to macroeconomic phenomena (e.g. Dong et al. 2020; Zhang, Gao, and Cai 2020; Vidal-Tomás and Alfaro 2020). Our paper contributes to the same literature and the proposed approach provides a visual laboratory to analyze risk.

Notes

1. All data (prices as well as other company-level information) have been collected from Thomson Reuters Eikon database. See Tables A4 in the Appendix for stock names and ticker symbols.
2. We note that this step can be conducted in multiple ways. Notably, Diebold and Yilmaz (2015a) have 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.
3. See Table 1 in the Online Supplementary Material for ARCH-LM test on the return series for all the stocks and ADF test results for the conditional volatility series. For the GARCH estimation, we have used *sgarch* module in *rugarch* package in R.
4. For a given network $G = (V, E)$ with adjacency matrix A , we define eigenvector centrality to be a vector c^{eig} which solves $Ac^{eig} = \lambda c^{eig}$ where λ is chosen to be the maximum eigenvalue λ_{\max} of the adjacency matrix A . For the purpose of finding eigenvector centrality on the correlation matrix, we consider $A = \Gamma$.
5. There are other centrality measures like degree, betweenness and closeness, etc. Here, we do not utilize them for three reasons. One, they are more useful for binary networks than weighted networks. Two, eigenvector centrality has algebraic properties that these measures do not possess. We have utilized these properties throughout the analysis. Three, eigenvector centrality has an interpretation in terms of market mode due to principal component analysis (see Section 5). Other centrality measures do not have this nice property.
6. Here y is a generic variable. For actual estimation, we substitute y_{it} by $\log(\sigma_{it})$. Taking $\langle roman \rangle \log \langle /roman \rangle$ is important to take care of the skewness of the latent volatility process.
7. Here we note two points. First, since we are using MST only for the purpose of visualization (all computation and estimation have been carried out on the whole network), our results do not depend in any way on the MST. Second, there could be other approaches to reduce the network size. However, MST by construction retains exactly $N-1$ edges for N nodes and still retain the connectivity of the network. Therefore, MST would produce the smallest non-trivial network. Since for visualization, our goal is to present as clean a picture as possible, we have implemented the MST.
8. 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). Interested readers can refer to Kreyszig (2015) for a textbook description of the minimum spanning tree and the relevant methodology used here.
9. In this paper, we have analyzed daily data.
10. Full network with all connections is too dense to visually identify the nodes, edges and how shocks are propagating.
11. Complete description of the company names and the corresponding short forms (ticker symbols) can be found in the Appendix. See Table A4 for the listed companies considered in this paper.
12. We thank a reviewer for this suggestion. See Figure 1 in the Online Supplementary Material for a frequency distribution of all the differences in magnitude between the estimated impulse response functions and their simulated counterparts. The differences are minimal.
13. See Tables 2 and 3 in the Online Supplementary Material for a rolling window estimation with yearly overlaps from 2002–05 to 2014–17.
14. *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.

Acknowledgments

This research was partially supported by institute grant, IIM Ahmedabad. We wish to thank conference and seminar participants at the WEHIA'19, ICU Tokyo, Reserve Bank of India (Strategic Research Unit), IIT Kanpur, IIT Delhi, Dibrugarh University, IIM Calcutta, two anonymous reviewers and Giuseppe Brandi for useful comments. All remaining errors are ours.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Indian Institute of Management Ahmedabad [R&P Grant].

References

- Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi. 2012. "The Network Origin of Aggregate Fluctuations." *Econometrica* 80: 1977–2016.
- Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi. 2015. "Systemic Risk and Stability in Financial Networks." *American Economic Review* 105 (2): 564–608.
- Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi. 2016. "Networks, shocks, and systemic risk." In *Handbook of Network Economics*, ed. 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 M. Hallin. 2017a. "Generalized Dynamic Factor Models and Volatilities: Estimation and Forecasting." *Journal of Econometrics* 201 (2): 307–321.
- Barigozzi, M., and M. Hallin. 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 C. Klymko. 2013. "A Matrix Analysis of Different Centrality Measures." *arXiv preprint arXiv:1312.6722*, p. 37.
- Bonacich, P. 1987. "Power and Centrality: A Family of Measures." *American Journal of Sociology* 92 (5): 1170–1182.
- Bonanno, G., G. Caldarelli, F. Lillo, S. Micciché, N. Vandewalle, and R. N. Mantegna. 2004. "Networks of Equities in Financial Markets." *European Physical Journal B* 38: 363–371.
- Bouchaud, J.-P., and M. Potters. 2003. *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*. New York: Cambridge university press.
- Cohen, L., and A. Frazzini. 2008. "Economic Links and Predictable Returns." *The Journal of Finance* 63 (4): 1977–2011.
- Corsi, F., F. Lillo, D. Pirino, and L. Trapin. 2018. "Measuring the Propagation of Financial Distress with Granger–Causality Tail Risk Networks." *Journal of Financial Stability* 38: 18–36.
- Demirer, M., F. X. Diebold, L. Liu, and K. Yilmaz. 2017. "Estimating Global Bank Network Connectedness." *Journal of Applied Econometrics* 33 (1): 1–15.
- Dessaint, O., T. Foucault, L. Frésard, and A. Matray. 2018. "Noisy Stock Prices and Corporate Investment." *Review of Financial Studies (forthcoming)* 32 (7): 2625–2672.
- Diebold, F. X., and K. Yilmaz. 2008. "Macroeconomic Volatility and Stock Market Volatility, Worldwide." Technical report, National Bureau of Economic Research.
- Diebold, F. X., and K. Yilmaz. 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 K. Yilmaz. 2015a. *Financial and Macroeconomic Connectedness: A Network Approach to Measurement and Monitoring*. New York: Oxford University Press.
- Diebold, F. X., and K. Yilmaz. 2015b. "Trans-Atlantic Equity Volatility Connectedness: US and European Financial Institutions, 2004–2014." *Journal of Financial Econometrics* 14 (1): 81–127.
- Dong, Z., H. An, S. Liu, Z. Li, and M. Yuan. 2020. "Research on the Time-Varying Network Structure Evolution of the Stock Indices of the BRICS Countries Based on Fluctuation Correlation, 2004–2014." *International Review of Economics & Finance* 69: 63–74.
- Engle, R., and S. Figlewski. 2014. "Modeling the Dynamics of Correlations Among Implied Volatilities." *Review of Finance* 19 (3): 991–1018.
- Fama, E. F., and K. R. French. 1992. "The Cross-Section of Expected Stock Returns." *Journal of Finance* 47 (2): 427–465.
- Foucault, T., and L. Fresard. 2014. "Learning From Peers' Stock Prices and Corporate Investment." *Journal of Financial Economics* 111 (3): 554–577.
- Garman, M. B., and M. J. Klass. 1980. "On the Estimation of Security Price Volatilities From Historical Data." *Journal of Business* 53: 67–78.
- Geraci, M. V., and J. -Y. Gnabo. 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., R. Morck, J. Shen, and B. Yeung. 2015. "Information, Analysts, and Stock Return Comovement." *The Review of Financial Studies* 28 (11): 3153–3187.
- Harvey, C. R., Y. Liu, and H. Zhu. 2016. "... and The Cross-Section of Expected Returns." *The Review of Financial Studies* 29 (1): 5–68.

- Hsu, N.-J., H.-L. Hung, and Y.-M. Chang. 2008. "Subset Selection for Vector Autoregressive Processes Using Lasso." *Computational Statistics & Data Analysis* 52 (7): 3645–3657.
- Hué, S., Y. Lucotte, and S. Tokpavi. 2019. "Measuring Network Systemic Risk Contributions: A Leave-One-Out Approach." *Journal of Economic Dynamics and Control* 100: 86–114.
- Hüser, A.-C., G. Halaj, C. Kok, C. Perales, and A. van der Kraaij. 2018. "The Systemic Implications of Bail-In: A Multi-Layered Network Approach." *Journal of Financial Stability* 38: 81–97.
- Kilian, Lutz, and Helmut Lutkepohl. 2017. *Structural Vector Autoregressive Analysis*. 1st ed. Cambridge: Cambridge University Press.
- Kreyszig, E. 2015. *Advanced Engineering Mathematics*. 10th ed. Hoboken: Wiley.
- Kumar, S., T. Di Matteo, and A. S. Chakrabarti. 2020. "Disentangling Shock Diffusion on Complex Networks: Identification Through Graph Planarity." *Journal of Complex Networks* 8 (3): cnaa023.
- 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 L. A. Pastur. 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." Hearing, United States Senate Committee on Banking, Housing, and Urban Affairs 2: 25.
- Novy-Marx, R.. 2013. "The Other Side of Value: The Gross Profitability Premium." *Journal of Financial Economics* 108 (1): 1–28.
- Plerou, V., P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley. 2002. "Random Matrix Approach to Cross Correlations in Financial Data." *Physical Review E* 65: 066126.
- Plosser, C. I. 2009, March. Redesigning financial system regulation. In a speech at the New York University Conference "Restoring Financial Stability: How to Repair a Failed System".
- Pozzi, F., T. Di Matteo, and T. Aste. 2013. "Spread of Risk Across Financial Markets: Better to Invest in The Peripheries." *Scientific Reports* 3: 1665.
- Sharma, K., B. Gopalakrishnan, A. S. Chakrabarti, and A. Chakraborti. 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)* 58 (1): 267–288.
- Tumminello, M., T. Di Matteo, T. Aste, and R. N. Mantegna. 2007. "Correlation Based Networks of Equity Returns Sampled At Different Time Horizons." *The European Physical Journal B* 55 (2): 209–217.
- Tumminello, M., F. Lillo, and R. N. Mantegna. 2010. "Correlation, Hierarchies, and Networks in Financial Markets." *Journal of Economic Behavior & Organization* 75 (1): 40–58.
- Vassalou, M., and Y. Xing. 2004. "Default Risk in Equity Returns." *The Journal of Finance* 59 (2): 831–868.
- Vidal-Tomás, D., and S. Alfarano. 2020. "An Agent-Based Early Warning Indicator for Financial Market Instability." *Journal of Economic Interaction and Coordination* 15 (1): 49–87.
- Zou, H.. 2006. "The Adaptive Lasso and Its Oracle Properties." *Journal of the American Statistical Association* 101 (476): 1418–1429.
- Zhang, J., Y. Gao, and S. Cai. 2020. "The Hierarchical Structure of Stock Market in Times of Global Financial Crisis." *Physica A* 542: 123542.

Appendices

In this section, we provide a list of all stocks analyzed and additional figures.

Appendix 1. Identity of the stocks

Ticker symbols of all stocks considered along with full company names are given in Table A4. Below we provide the list of all ticker symbols.

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.

Appendix 2. Additional network figures

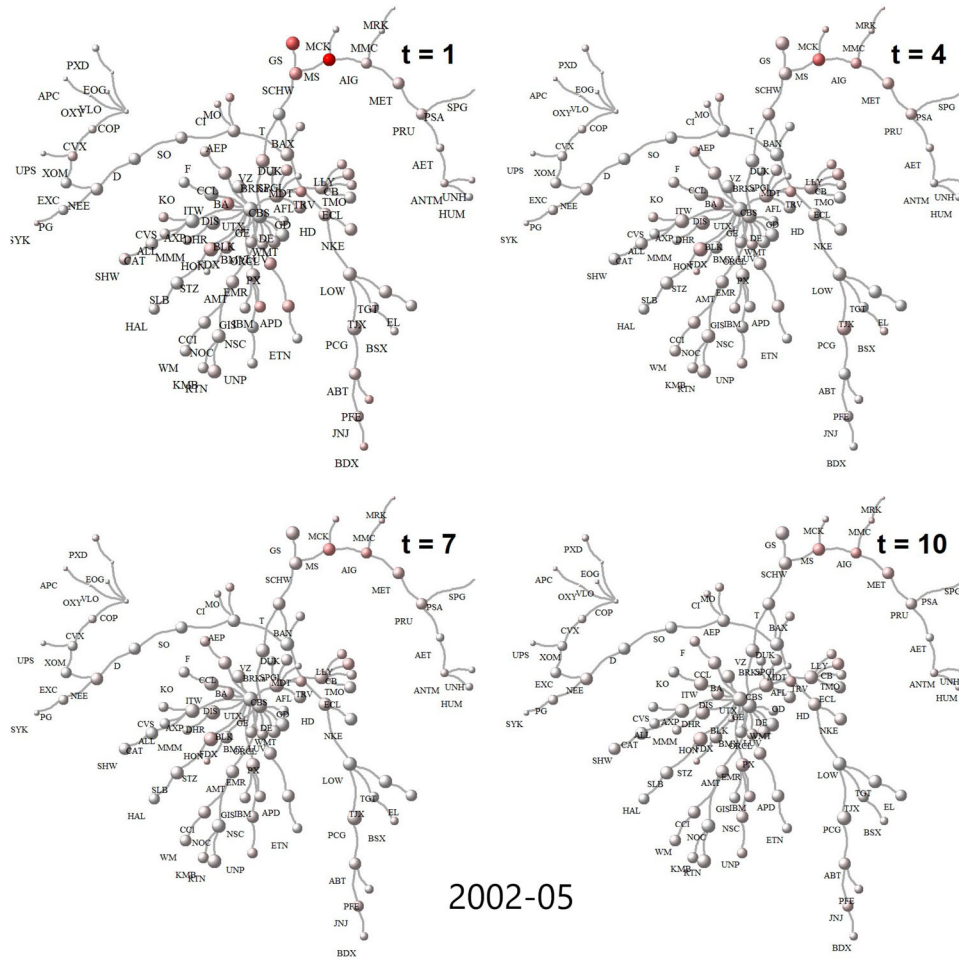


Figure A1. Ripple effects over 10 days ($t = 1$ (left-top panel); 4 (right-top panel); 7 (left-bottom panel); 10 (right-bottom panel)) on the network of top 100 firms in terms of market capitalization during the period 2002–05. The epicenter of the shock is Morgan Stanley (MS). Intensity of color represents the magnitude of return volatility of the nodes in the network.



Figure A2. Ripple effects over 10 days ($t = 1$ (left-top panel); 4 (right-top panel); 7 (left-bottom panel); 10 (right-bottom panel)) on the network of top 100 firms in terms of market capitalization during the period 2010–13. The epicenter of the shock is Morgan Stanley (MS). Intensity of color represents the magnitude of return volatility of the nodes in the network.

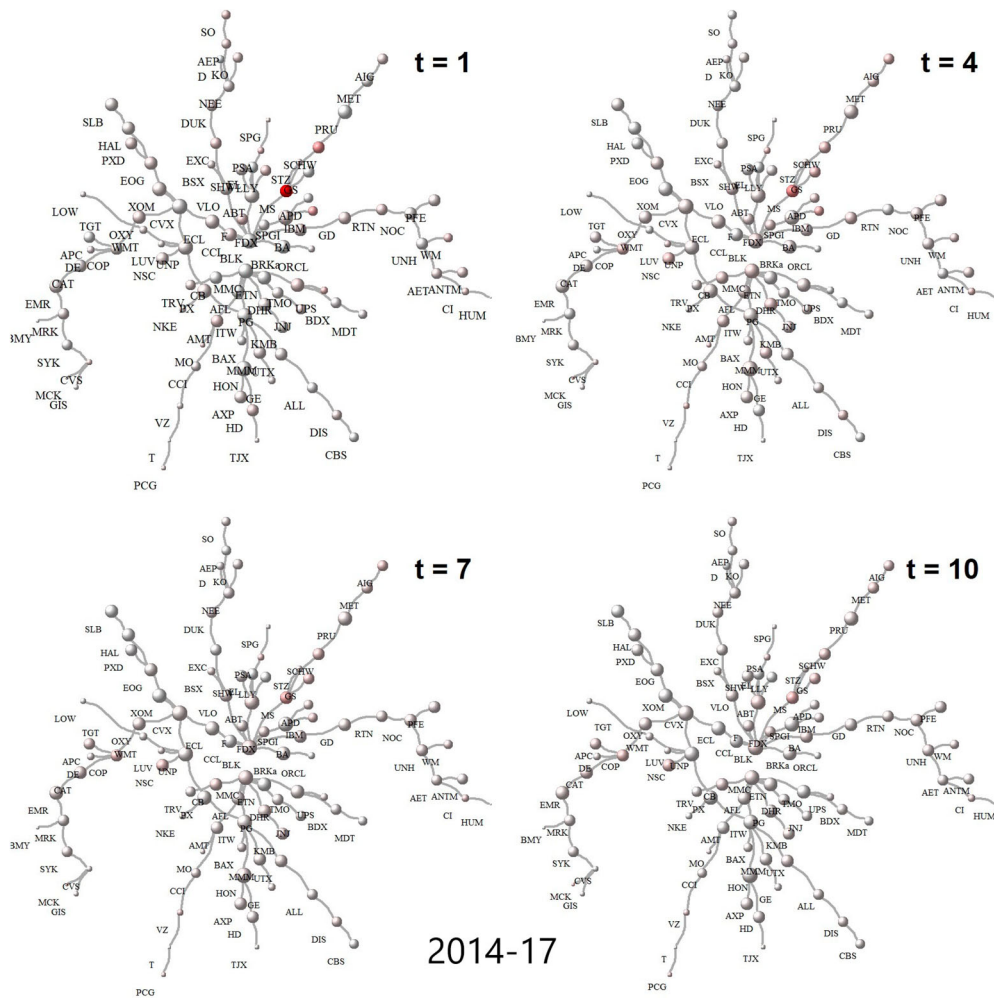


Figure A3. Ripple effects over 10 days ($t = 1$ (left-top panel); 4 (right-top panel); 7 (left-bottom panel); 10 (right-bottom panel)) on the network of top 100 firms in terms of market capitalization during the period 2014–17. The epicenter of the shock is Morgan Stanley (MS). Intensity of color represents the magnitude of return volatility of the nodes in the network.

Appendix 3. Tables

Table A1. Rank correlation matrix of return centrality vectors c_r^{eig} across time.

	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)	0.64	1	–	–
c_r^{eig} (10–13)	0.51	0.57	1	–
c_r^{eig} (14–17)	0.54	0.42	0.70	1

Description: The vector c_r^{eig} represents the eigenvector centralities of stocks in the return correlation matrix. For each snapshot, there exists one unique vector and the stocks can be ordered with respect to their rank in terms of centrality. The above table represents the (Spearman) correlation between the rank-ordering across four snapshots in time (2002–05, 2006–09, 2010–13 and 2014–17).

Table A2. Rank correlation matrix of volatility centrality vectors c_v^{eig} across time.

	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)	0.33	1	–	–
c_v^{eig} (10–13)	0.17	0.49	1	–
c_v^{eig} (14–17)	0.01	0.32	0.50	1

Description: The vector c_v^{eig} represents the eigenvector centralities of stocks in the latent volatility correlation matrix. For each snapshot, there exists one unique vector and the stocks can be ordered with respect to their rank in terms of centrality. The above table represents the (Spearman) correlation between the rank-ordering across four snapshots in time (2002–05, 2006–09, 2010–13 and 2014–17).

Table A3. Cross section (OLS) and panel (random effect) regression results.

	Dependent variable: return centrality										
	OLS								panel: Random linear		
	2002–05		2006–09		2009–13		2014–17		(Individual)	(Time)	(Two-ways)
Logasset	0.279*** (0.075)	0.262*** (0.088)	0.057 (0.100)	0.050 (0.092)	0.222*** (0.081)	0.221*** (0.084)	0.179** (0.084)	0.210** (0.084)	0.195*** (0.058)	0.195*** (0.043)	0.195** (0.086)
ROA		0.003 (0.018)		0.014 (0.012)		0.012 (0.018)		0.014 (0.017)	0.015** (0.007)	0.015* (0.008)	0.015 (0.011)
Risk rating		−0.001 (0.005)		−0.007 (0.005)		−0.009** (0.003)		−0.003 (0.004)	−0.004** (0.002)	−0.006** (0.002)	−0.004 (0.003)
Analysts		−0.006 (0.014)		0.041* (0.022)		0.002 (0.013)		−0.019* (0.011)	−0.011 (0.007)	−0.003 (0.007)	−0.011 (0.011)
Constant	−6.729*** (1.837)	−6.224*** (2.153)	−1.395 (2.439)	−1.945 (2.166)	−5.488*** (2.000)	−5.474** (2.173)	−4.438** (2.086)	−4.790** (2.218)	−4.582*** (1.404)	−4.741*** (1.068)	−4.584** (2.086)
Observations	100	96	100	98	100	98	100	100	392	392	392
R ²	0.134	0.116	0.005	0.079	0.063	0.126	0.038	0.081	0.046	0.070	0.065
Adjusted R ²	0.125	0.077	−0.006	0.039	0.054	0.088	0.028	0.043	0.036	0.060	0.055
Res. Std. Error	0.935	0.925	1.003	0.976	0.973	0.956	0.986	0.978			
F Statistic	15.148***	2.980**	0.449	1.991	6.637**	3.353**	3.898*	2.103*	4.672***	7.243***	6.718***
df	1; 98	4; 91	1; 98	4; 93	1; 98	4; 93	1; 98	df = 4; 95	4; 387	4; 387	4; 387)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Description: This table represents regression results with C_r^{eig} as the dependent variable and firm-level fundamentals (size captured by log of total assets, profitability captured by return on assets, financial health captured by risk rating and popularity captured by number of analysts). Firms' total assets clearly correlates with return centrality. The panel regression in the end is consistent with the OLS results.

Table A4. Ticker symbols along with company names of top 100 stocks in terms of market capitalization in 2017.

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