On the Consequences of Stock Network Topology on Portfolio Diversification

Abalfazl Zareei[a], Gustavo Peralta[b],

Abstract

This paper investigates the impact of stocks network topologies on the portfolio risk. We construct stocks directed network using the lead-lag relationships among stocks obtained from Vector Autoregressive (VAR) model. The portfolio variance of a naïve diversification rule and its asymptotic behavior for several stylized network structures are analyzed. We conclude that in an homogeneous and symmetric network, portfolio variance is lower than in a network structure where a few stocks assume more central positions. Moreover, we find that the impact of shocks in a centrality placed stocks of a star–like network on the short-term return-to-risk ratio is higher and longer-lasting compared to a more homogeneous structure. When analyzing long-term impact of shocks, we relate the return of portfolio to the Bonacich centrality of stocks in the portfolio's network structure. Thereby, the higher is the centrality of the stock in the network structure, the higher would be the impact on portfolio from a shock to this stock.

Keywords: *Diversification, Network Theory, Vector Autoregressive Model*

a Abalfazl Zareei is at the Department of Business Administration, Universidad Carlos III de Madrid, c/ Madrid 126, 28903 Getafe (Madrid, Spain) Email: azareei@emp.uc3m.es

b Gustavo Peralta is at the Department of Research and Statistics at CNMV and Department of Business Administration at Universidad Carlos III de Madrid. Email: gperalta@cnmv.es

1. Introduction

A well know proverb in finance states *"do not put all of your eggs in the same basket"* as a way to stress the benefit of diversification. The standard theory shows a decreasing and monotonic relationship between portfolio variance and the number of stocks in a portfolio. This negative relationship remains until the portfolio variance reaches to its asymptotic lower limits given by the mean correlation of returns, commonly associated to the systematic risk (Marín & Rubio 2011). In this paper, we specifically study how particular patterns of interconnection between stocks in a network structure affect the portfolio diversification benefits.

The contribution of this paper is three-fold. First, we show theoretically how the dynamic interaction of stocks obtained from a Vector Autoregressive (*VAR*) model effect the portfolio variance. Second, we analyze how various stylized network structures would result in different portfolio variance and how this variance converges as we increase the number of stocks in these structures. The diversification rule in star-like network structures are found to digress from other type of structures signifying how asymmetry in the number of out-going links in a stocks network raises the portfolio variance. Third, by investigating the impact of shocks (where? In the center?) on short-term return-to-risk ratio in different network structures, we show that star-like network structures are more affected and these impact are found to be longer-lasting.

We assume a dynamic structure for the generative return's process captured by a Vector Autoregressive specification (VAR). Such specification allows us to account for lead-lag relationships between stock's returns that we interpret as directed links in a network structure. To be more concrete, our *directed and weighted stock network* is composed of nodes representing stocks and the adjacency matrix is given by the coefficients matrix from the *VAR(1)* model. Thereby, a non-null and significant coefficient accounting for effect of stock *j* in the return's process of stock *i* implies the existence of a link from node *j* to node *i* in the stock network. Our sample comprises return data for S&P500 starting from March 2004 and finishing in April 2013, thus covering pre-crisis, crisis and post-crisis periods. Since sparsity in the adjacency matrix is required for interpretability of the network approach, the VAR model is estimated using Elastic.Net algorithm. ((Zou & Hastie 2005) and (Zou & Zhang 2009))

The paper is divided into a theoretical and an empirical part. In the theoretical part, we provide a theoretical proof of the interactions between the portfolio variance and the stock network's structure by means of stylized architectures. Under this setting, we prove that the asymptotic limit of portfolio variance is equal for a disconnected network, fully connected network and a circle network[1]. However, as long as some stocks assume dominant positions into the network, like in a start network, the variance of the portfolio does not converge toward its presumed minimum but remains higher. More generally, for those network structures where the role taken by different stocks tend to be symmetric and homogeneous, the diversification argument remains valid. This result is attenuated when a few firms in the market takes more central positions since such lower bound is never reached. Thus, higher concentration in the out-degree distribution in the network structure prevents reaching the maximum diversification benefits. Additionally, we find the long-term impact of shocks on the central stocks (ranked by Bonacich centrality) to be more influencing the portfolio return comparing to low central stocks.

The second part of the paper presents an empirical assessment of the stock network. Since the recent financial crisis took place approximately in the middle of our sample period, we characterize the state of the stock network in the prior, during and post crisis periods. We find that before and after crisis, the stocks network has more resemblance to the star network. Additionally, during the crisis period, the network structure becomes denser signifying higher number of interactions between stocks in concordance with the findings in Billio et al. (2012).

From our point of view, the question addressed in this article is fundamental in order to get a deep understanding on the benefit of portfolio diversification. Our major proposition clearly states that the advantage of the portfolio diversification is constrained by the topology of the stock network. Our results are comparable with those in (Acemoglu et al. 2012) for the aggregate variance of the economy and thus embracing the idea that sometimes average measures uncovers the effects excreted by extremely influential nodes in the underlying network.

The remainder of the paper is organized as follows. Section 2 presents a literature review regarding the current study. Section 3 defines primary steps in constructing the directed network. In section 4, we investigate various special directed network cases. Section 5 presents the impact of shocks to the stocks in the stock network on the portfolio returns and diversification benefits. Section 6 establishes the main results and section 7, presents the implication of findings in this paper. Finally, section 8 concludes and outlines future research lines.

[1] In a disconnected network, each stock is affecting itself, and in a fully connected network each stock is connected with the rest of stocks in the market. In a circle network a node *i* is affected only by stock *i-1* and affects stock *i+1*.

2. Literature Review

The current paper is grounded in two different branches of literature. The first one corresponds to the study of stock markets through the lens of network theory. Second one covers the literature regarding the diversification. Next, we briefly describe their salient findings.

The study of stock markets by means of network theory could be subdivided into three subranches. The first subsection is closely related to physics and it is basically concerned with the study of the topology of the related network. (Mantegna 1999) and (Bonanno et al. 2001) were among the first to apply the so-called *Minimum Spanning Tree (MST)* for the US Market with empirical correlation matrix of stock returns to uncover corresponding stock network.[2] A similar approach was taken by (Vandewalle et al. 2001) also for the US market and by (Jung et al. 2006), (Garas & Argyrakis 2007) and (Huang et al. 2009) for the Korean, Greek and Chinese markets, respectively. Among their main results, it is interesting to highlight: *i)* how different branches of the stock network coincided to specific economics sectors, *ii)* power-law degree distributions and the correlation in the degrees of connected nodes are evidence of non-random arrangement of links. In (Onnela et al. 2003), a dynamic perspective of the same framework is undertaken allowing the authors to study the properties of the stock network through time. For authoritative summaries of the field see (Bonanno et al. 2004) and (Tumminello et al. 2010).

A second subgroup of network-related articles comes from the econometric and financial literature using network approach to get new insights about systemic risk issues. In (Billio et al. 2012), the authors build a directed Granger-causality network[3] to capture the market interconnectedness. In this structure the links account for statistically significant pairwise lead-lag relationships between institution's monthly returns. The concern about correlation in the tails of the return distribution results in the estimation of the so-called *tail-risk network* in (Hautsch et al. 2014a). This structure is a weighed-directed network in which the links between institutions are given by the interconnectedness of firm's Value-at-Risk. Once this network is in place, the authors compute the *realized systemic risk beta* corresponding to the systemic relevance of financial firms in the market. In a

[2] MST is a filtering technique that allows us to build a connected network of N stocks by joining together pairs of them in accordance to their pair correlation (in decreasing order) as long as no loops is formed in the structure. The resulting network is a tree network.

[3] The authors also measure connectedness through principal components but this approach is not the focus of the current paper.

closely related paper (Hautsch et al. 2014b), the authors adapt the *tail-risk network* framework to account for forecasting purpose of firm's systemic relevance. Finally, (Diebold & Yılmaz 2014) measures the connectedness in a network in which the links between financial institutions are assigned in accordance to the variance decomposition of the volatility forecast error, giving rise to a volatility weighed-directed network. They show how the cycles of the total connectivity in the structure coincide with major disruptions in the US market.

The final subgroup of network-based articles argues in favor of the network approach as a promising tool to enhance portfolio's performances. (Pozzi et al. 2013) shows the improvements in financial performance of an investment strategy that assigns wealth toward stocks belonging to the periphery of stock network. This unconditional strategy is put into question in (Peralta and Zareei 2014). In this paper, the authors propose a dynamic strategy, the so-called *ρ-dependet strategy*, accounting not only for the position of different stocks in the market but also for their performance in isolation. A final paper worth to mention is (Ozsoylev et al. 2014) where the informational diffusion process was studied for the Istanbul Stock Exchange. The major contribution of this article regards on the quantification of the early trading advantages and higher returns obtained by investors centrally placed in such informational network.

The literature regarding financial diversification is enormous as the reader should expect. In what follows, the salient aspects are discussed. Among the earliest studies, (Samuelson 1967) clearly states the theoretical condition under which diversification among a fixed number of stocks pays off. He shows that positive diversification (a situation in which each assets enters in the portfolio with a positive weight) holds when stocks are independently distributed or when they present negative correlations. However, for the case of positive correlations, positive diversification does not necessarily holds. From a more empirical perspective, (Evans & Archer 1968) focus on the relationship between portfolio dispersion and the number of stocks in a randomly selected and equally weighted portfolio. They show a stable and predictable association which is characterized by a rapidly decreasing and asymptotic function, with the asymptote approximation of the systematic risk. They also raise some doubts about the economic justification of portfolio sizes beyond 10 stocks without a proper marginal cost-benefit analysis. (Mao 1970) provides theoretical support for this finding arguing that relatively few stocks are required to capture the bulk of the benefit of diversification (17 under reasonable conditions).

The paper of (Evans & Archer 1968) has raised many critics from the research community about the number of stocks needed to obtain the maximum level of diversification. Among the first group of critics (Elton & Gruber 1977) and (Bird & Tippett 1986) argues that the parametric relationship in (Evans & Archer 1968) is misspecified. Both articles provide the exact equations connecting portfolio risk and its size showing that there is still room for risk reduction well beyond the 10 stock proposed by (Evans & Archer 1968). A second group of critics focus on the nature of the systematic risk and particularly on the beta coefficient. The so-called beta effect, say the positive association between portfolio systematic risk (beta) and idiosyncratic risk (calculated as the variance of the residuals from the market model), is documented in (Klemkosky & Martin 1975). Under such condition, high beta portfolios requires larger number of securities to achieve approximately the same level of diversification than low beta ones. The non-stationarity of the beta coefficient is considered in (Chen & Keown 1981) by providing efficient estimates for both, systematic and idiosyncratic risk under consideration. In a similar fashion (Campbell et al. 2001) reports a decreasing correlation between the returns of individual stocks and the market. This phenomena results in an increment in the idiosyncratic risk and a rise in the number of stocks needed to obtain a given level of portfolio diversification. This evidence is not consistent with the more recent data including the last financial crisis showing an increase in the market correlation among security types and among international markets (James et al. 2012).

It is worth mentioning two more studies embarrassing a different methodology to assess the benefit of diversification. (Statman 1987) relies in a Security Market Line approach finding that well-diversified portfolios must include at least 30 stocks and at least 40 if no leverage is employed. The shortfall risk approach is put into practice in (Domian et al. 2007). Two main conclusions come out from this study are: *i)* for long-term investors who wish to outperform Treasure bonds, a portfolio with more than 150 stocks is required and *ii)* greater risk reduction is archived by increasing such number than by spreading the holdings across industries.

3. Formation of Directed Stock Network

In general terms, a network is a pair of sets $\varphi = \{V, E\}$, where $V = \{1, 2, \dots, n\}$ is the set of nodes and E the set of links connecting pairs of nodes. If there is a link from node *i* to node *j*, $(i, j) \in E$. A convenient way to arrange the information contained in E is by means of the so-called adjacency matrix $A_{n\times n} = [A_{ij}]_{n\times n}$. When $A_{ij} \neq 0$, there is a relationship between node *i* and node *j*. The network is said to be undirected if $A = A^T$, therefore $(i, j) \in E$ also implies $(j, i) \in E$. Note that for undirected network no-causal relationship is attached to the links and they are visually represented as a line, $(j - i)$. On the other hand, if $A \neq A^T$, the network is said to be directed and A_{ij} entails a causal relationship from node j to node i which does not necessarily implies the reverse. In this case, the links are visually represented as arrows, $(j \rightarrow i)$. Furthermore, if $A_{ij} \in \{0,1\}$, φ is said to be an un-weighted network and when $A_{ij} \in \mathbb{R}$, each link in the network carries information about interaction intensity between nodes leading to a weighted network[4].

Suppose an investor selects a basket of n stocks whose returns follow a Gaussian Vector Autoregressive model of order 1, *VAR(1),* as in equation (1).

$$r_{t+1} = a + Br_t + u_{t+1} \tag{1}$$

Where r_t is the n dimensional vector of stock returns in period *t*, a is the vector of intercepts allowing for non-zero expected returns and $B = [b_{ij}]$ is an nxn matrix where the element b_{ij} represents the impact of stock *j* in period *t* on stock *i* in period *t+1*. Finally, u_{t+1} is a Gaussian white noise process with zero mean vector and positive definite covariance matrix $\Sigma_u = [\sigma_{u,ij}]$.

VAR model has been used before to capture the dynamical dependency between stocks - See DeMiguel et al. (2014), Eum and Shim (1989) and Chordia and Swaminathan (2000)[5]. DeMiguel et al. (2014) exploit the individual stock returns serial dependency via *VAR* model for the purpose of achieving better out-of-sample performance. They develop arbitrage and mean-variance strategies based on *VAR* model and find these strategies to perform well when the transaction costs are below ten basis points. Eum and Shim (1989) apply *VAR* model to daily return of nine market indices in order to investigate the interactions between stock markets. Chordia and Swaminathan (2000), investigating the impact of trading volume on lead-lag patterns in stock returns, apply *VAR* model to two portfolios, one with high-trading-volume stocks and other with low-trading-volume stocks. In an informational efficient market, there should not exist any dynamical pattern between stock returns, however, as it is stated in Billio et al (2012), in the presence of market frictions (e.g. value-at-risk constraints, transaction costs, borrowing constraints, costs of gathering and processing

[4] The reader is referred to (Newman 2010) and (Jackson 2010) for a comprehensive treatment of the field.

[5] Particularly, in strategic asset allocation, VAR models proved to be a useful tool. We refer to Campbell and Viceira (1999, 2002), Campbell, Chan and Viceira (2003), and Barberis (2000).

information, institutional restrictions on shortsales), we may find causal relations among returns, even though not necessarily exploitable.

We define two different but related stock market networks, $\Phi^W = \{N, B\}$ and $\Phi^U = \{N, E\}$. In both cases, N accounts for the set of stocks in the investor's portfolio. Φ^W is a directed and weighted network whose adjacency matrix is given by B. Thereby, a weighted link from stock *j* to stock *i* exists as long as b_{ij} is different from zero. Φ^U is the un-weighted version of Φ^W with adjacency matrix $E = [e_{ij}]$ determined as follows:

$$e_{ij} = \begin{cases} 1, & b_{ij} \neq 0 \\ 0, & b_{ij} = 0 \end{cases} \tag{2}$$

As an example, in figure 1 we provide the network of 100 highly capitalized stocks listed in NYSE in the period 1965 to 2006[6]. In this network, the nodes are the stocks and the links accounts to the pairwise lead-lag relationship among them. By construction, those links are directed giving rise to type of links: those pointing toward a node (in-links) and those pointing from that node (out-links). The darker the node, the higher is its out-degree.

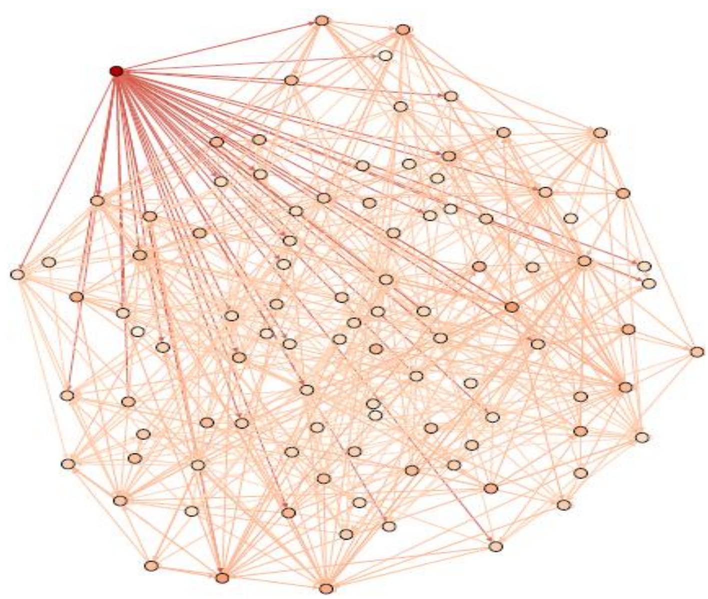

Figure 1. Directed network for open-close daily returns of 100 highly capitalized stocks from 1995 to 2006

[6] For a detailed explanation of the *VAR* estimation, the reader is referred to section 6.

4. Portfolio diversification and network topology

We assume that the return process in (1) is stationary an as a consequence, each eigenvalue of B has modulus less than 1. Under this condition, it is said that the process is well-behaved with mean vector and covariance matrix given as follows[7]:

$$E(r_t) = \mu_r = (I_N - B)^{-1}a \quad (3)$$

$$\Sigma_r = \sum_{i=0}^{\infty} B^i \Sigma_u B^{i'} \quad (4)$$

Due to the imposition of stationarity, B^i converges to zero rapidly with increasing *i*.[8] Therefore, a convenient approximation of equation (4) is given by equation (5) where the term $B^i \Sigma_u B^{i'}$is negligible for $i \geq 2$

$$\tilde{\Sigma}_r = \Sigma_u + B\Sigma_u B' \quad (5)$$

The variance of a portfolio composes by each of the stocks in N with weight vector, $\boldsymbol{w}$, is as follows:

$$\sigma_p^2 = \boldsymbol{w}'\Sigma_u \boldsymbol{w} + \boldsymbol{w}' B\Sigma_u B' \boldsymbol{w} \quad (6)$$

$$\sigma_p^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{u,ij} + \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j (\sum_{k=1}^{n}\sum_{l=1}^{n} B_{ik} B_{jl} \sigma_{u,kl}) \quad (7)$$

The right-hand side of equation (7) is composed of two terms, the first one corresponding to the traditional portfolio variance and the second one associated to the dynamical part of the return process which also captures the impact of the network topology $\boldsymbol{\Phi^W}$. Specifically, the $\boldsymbol{\Sigma_u}$ in the *VAR* model is not diagonal. It represents the unconditional covariance between returns and

[7] We refer to (Lütkepohl 2007) for a detailed explanation.
[8] In order to demonstrate how fast it converges to zero, we provide a numerical example in appendix A.

thereby, we can verify the traditional notion of systematic and idiosyncratic risk of the portfolio from the first term in equation (7).

In order to verify how different patterns in the network structure impact the portfolio variance, figure 2 demonstrates the full list of network motifs (specific patterns of interactions) that fundamentally determines σ_p^2.

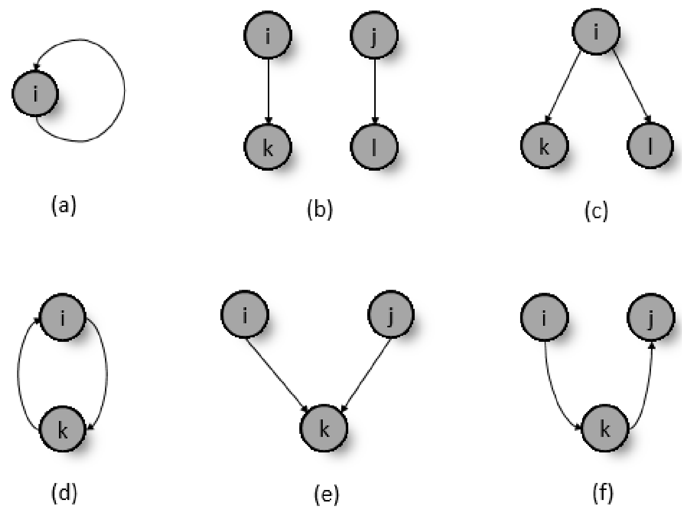

Figure 2. Specific network motifs influencing portfolio variance

The detailed explanation of the impact of each motif depicted in figure 2 on the portfolio variance is given in Appendix B. In total, the weight invested in the stocks with in-going links (vulnerable stocks), and the individual variance and covariance between the stocks with out-going links (threatening stocks) are the major determinant of portfolio variance. This is predictable as any variation in a vulnerable stocks comes from the threat exerted by threatening stocks.

4.1. Portfolio diversification and special cases of network topology

In this section, we consider various special network topologies and investigate the portfolio variance under these structures. Throughout this section, we impose simplifying assumptions to make our analysis more tractable. We consider a naïve investor that allocates his wealth equally among assets in the investor's portfolio set. As a consequence, the (column) vector of portfolio's weights is $\boldsymbol{w} = \frac{1}{n}\mathbf{1}$ where $\mathbf{1}$ is a column vector of ones. Additionally, with the aim to isolate any other possible effect different from the stock network topology, it assumed that $\sigma_{u,ij} = \sigma_u^2$ for $i = j$ and $\sigma_{u,ij} = \rho\sigma_u^2$ for $i \neq j$ where ρ accounts for the equal pair correlation of return. Finally, $b_{ij} = b < 1$ for $\forall ij$. When all these sets of simplifying assumptions are in place for a particular formula, we say that formula is under *SSA*.

When there is no dynamical structure in the return process, B is equal to the null matrix and as a consequence the expression (7) for the portfolio variances lead to its traditional formulation (Markowitz 1952). Therefore,

$$\sigma_p^2 = \boldsymbol{w}'\Sigma_u\boldsymbol{w} = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{u,ij} \tag{8}$$

Consistent with (Mao 1970), equation (8) under *SSA* is written as in (9) stressing the fact that portfolio variance is a function of its size.

$$\sigma_p^2(n) = \left[\frac{1}{n} + \rho\left(1 - \frac{1}{n}\right)\right]\sigma_u^2 \tag{9}$$

Let us consider two extreme cases, *i)* none diversification where $n = 1$ leading to $\sigma_p^2(1) = \sigma_u^2$ and *ii)* extreme diversification in which $n = \infty$ (when every stock in the market is included into the investment basket) where $\sigma_p^2(\infty) = \rho\sigma_u^2$. Note that since $\sigma_p^2(1) > \sigma_p^2(\infty)$ there exist diversification benefit.

When there is a dynamical structure, the variance of portfolio, in comparison with the no-dynamical case (see equation (8)), is increased by the element $\boldsymbol{w}'B\Sigma_u B'\boldsymbol{w}$ which in turn captures the effect of the stock network topology Φ^W. Next, we assume that the set of stocks included in the investment opportunity set form stylized network topologies that are carefully selected to gain new insights about their impact upon portfolio variance.

***Case I**: Disconnected Network*

The first case under analysis assumes each stock is affected only by itself as it is depicted in figure 3 and thus there is no dynamic interaction beyond their own autocorrelation process.

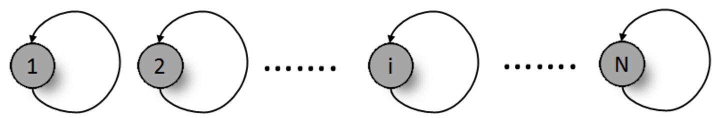

Figure 3. Special case where each stock is threatening only itself

In this case, the matrix B is diagonal and since there is not cross interaction $b_{ij} = 0 \; for \; i \neq j$; therefore the variance of portfolio is as follows:

$$\sigma_p^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{u,ij} + (\sum_{i=1}^{n} (w_i)^2 {b_{ii}}^2 {\sigma_{u,i}}^2 + \sum_{k=1}^{n}\sum_{l=1 \; k\neq l}^{n} w_k w_l b_{kk} b_{ll} \sigma_{u,lk}) \tag{10}$$

The first term of expression (9) is the usual portfolio variance and the second and third terms represent the impact from network topology. In the second and third terms, the weights, variance and covariance of each of the stocks in the investment set affect the portfolio variance. With regard to variance, the impact is positive for all of stocks. Under *SSA*, as long as $|b| < 1$, stationarity holds. In this case the variance of the portfolio is simplified to:

$$\sigma_p^2(n) = \left[\frac{1}{n} + \rho\left(1 - \frac{1}{n}\right)\right][1 + b^2]\sigma_u^2 \tag{11}$$

As before, two extreme cases regarding the value of n are mentioned. For $n = 1$, the portfolio variance is $\sigma_p^2(1) = [1 + b^2]\sigma_u^2$ and for the case of maximal diversification, $n = \infty$, $\sigma_p^2(\infty) = [1 + b^2]\rho\sigma_u^2$.

Case II: *Star Network*

Case (I) assumes an extreme situation of none interaction in the stock network. Another interesting is when there is one stock influencing all the other stocks in the portfolio.

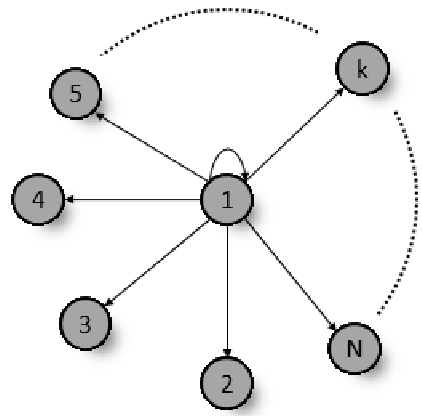

Figure 4. Star Network

In this case, B is a zero matrix expect for the column *i* which is composed by elements equal to b_{ij} for $i = 1,2,..,n$. Accordingly, the variance of the portfolio is given by expression (12). Considering the second term related to the network topology, we see that variance of the central stock, stock *1*, and also the weights allocated to stocks in the periphery of the network, w_i for $i = \{2,3,\ldots,n\}$, are relevant for determining σ_p^2.

$$\sigma_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{u,ij} + \sum_{k=2}^{N}\sum_{l=2}^{N} w_k w_l b_{k1} b_{l1} {\sigma_{u,1}}^2 \tag{12}$$

When *SSA* is imposed, matrix B has all its entries equal to zero except for the components of column *i* which are equal to b. For matrix B, all its eigenvalues are zero except for one which is equal to b. Therefore, as long as $|b| < 1$, stationarity holds. In such situation the portfolio is

$$\sigma_p^2(n) = \left[\frac{1}{n} + \rho\left(1 - \frac{1}{n}\right) + b^2\right]\sigma_u^2 \tag{13}$$

When $n = 1$, the return-to-risk ratio is $\sigma_p^2(1) = (1 + b^2)\sigma_u^2$ and for the case of maximal diversification characterized by $n = \infty$, $\sigma_p^2(\infty) = (\rho + b^2)\sigma_u^2$.

Case III: *Inverse Star Network*

The inverse of case II is presented in figure 5 in which the center of the network receives impacts from other stocks in the investment set.

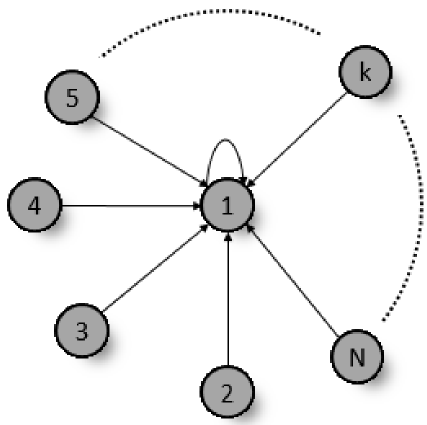

Figure 5. Inverse Star Network

In this case, B is a zero matrix expect for the row *i* which is composed of elements equal to b_{ij} for $i = 1,2,..,n$. The corresponding portfolio variance is given by equation (14). Focusing on the second term of equation (14), we see that the only weight involved in the formula is the one from the central stock. However, all of the correlations between pairs of peripheral stocks are crucial determinents in equation (14).

$$\sigma_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{u,ij} + (w_1)^2 \sum_{k=2}^{N}\sum_{l=2}^{N} b_{1k} b_{1l} \sigma_{u,kl} \qquad (14)$$

The structure of eigenvalues for matrix B is exactly the same as in the star network because these matrices are the transpose of each other. In this case, B is a zero matrix expect for the row *i* that is composed of elements equal to b. Therefore, as long as $|b| < 1$, stationarity holds. In this case, the variance of the portfolio is

$$\sigma_p^2(n) = \left[\frac{1}{n} + \rho\left(1 - \frac{1}{n}\right)\right][1 + b^2]\sigma_u^2 \qquad (15)$$

An interesting result comes when one compares expressions (14) and (11). It turns out that the portfolio variance is exactly the same in these two specifications; however, this is not true for the star network case. This highlights the importance of the concentration in the degrees pointing from a particular node (out-degree) in comparison to the concentration of the degrees pointing to a particular node (in-degree). It is clear that a highly out-degrees concentration in the stock network undermines the benefits of diversification.

***Case IV**: Circle Network*

Another interesting case to study regards to the circle network which is depicted in figure 6. In this symmetric structure stock i affects stock $i+1$ and it is being affected by stock $i-1$.

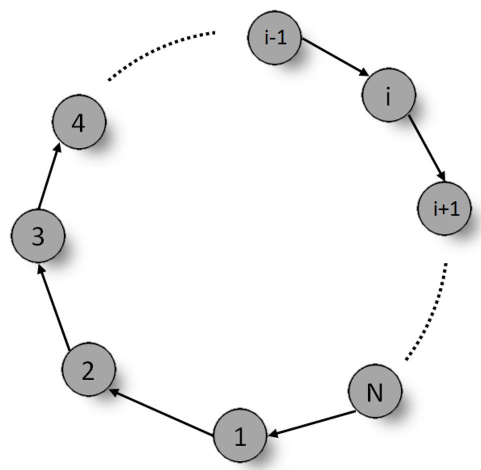

Figure 6. Circle Network

The elements of matrix B are zero except for those located in first diagonal below the main diagonal and for the one located in the upper right corner. The portfolio variance in this case is given by equation (16). Note that the variance of portfolio due to the network structure is determined by the weights and covariance of consecutive stocks.

$$\sigma_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \sigma_{u,ij} + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j b_{i(i+1)} b_{(j)(j+1)} \sigma_{u,(i+1)(j+1)} \qquad (16)$$

Under *SSA*, the statitionary is preserved as long as $|b| < 1$. We see that the variance is exactly equal to the one from the disconnected network.

$$\sigma_p^2(n) = \left[\frac{1}{n} + \rho\left(1 - \frac{1}{n}\right)\right][1 + b^2]\sigma_u^2 \tag{17}$$

Case V: *Fully Connected Network*

Finally, the case in which the network is fully connected is depicted in figure 7 for the case of $n = 4$. In this situation each stock is connected with the rest of the stocks in the investment set.

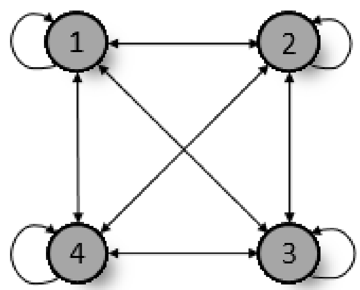

Figure 7. Fully Connected Network

This situation is somehow different since in order to preserve stationary under *SSA*, the parameter b must be sufficiently small in relation to n. Therefore, it is assumed that $b = \frac{1}{n}\delta$ for $0 < \delta < 1$. It could be proved that the eigenvalues of B are all equal to zero expect for the largest which is equal to δ. The portfolio variance is as follows:

$$\sigma_p^2(n) = \left[\frac{1}{n} + \rho\left(1 - \frac{1}{n}\right)\right][1 + \delta^2]\sigma_u^2 \tag{18}$$

As before, two extreme cases regarding n are mentioned. For $n = 1$, the portfolio variance is $\sigma_p^2(1) = [1 + \delta^2]\sigma_u^2$. Note that when $n = 1$, then $b = \delta$. For the case of maximal diversification characterized by $n = \infty$, $\sigma_p^2(\infty) = [1 + \delta^2]\rho\sigma_u^2$. The reader should note that expression (18) presents the same analytical structure than those prevailing for the disconnected network and it is identical as long as $b = \delta$.

4.2. Numerical experiment

Next, we investigate the convergence of portfolio variance for increasing number of stocks under different network topologies under consider SSAFor comparison purpose, figure 8 provides the behavior of the portfolio variance considering the cases of No-Dynamical, Disconnected Network and Star Network for different portfolio sizes. [9]

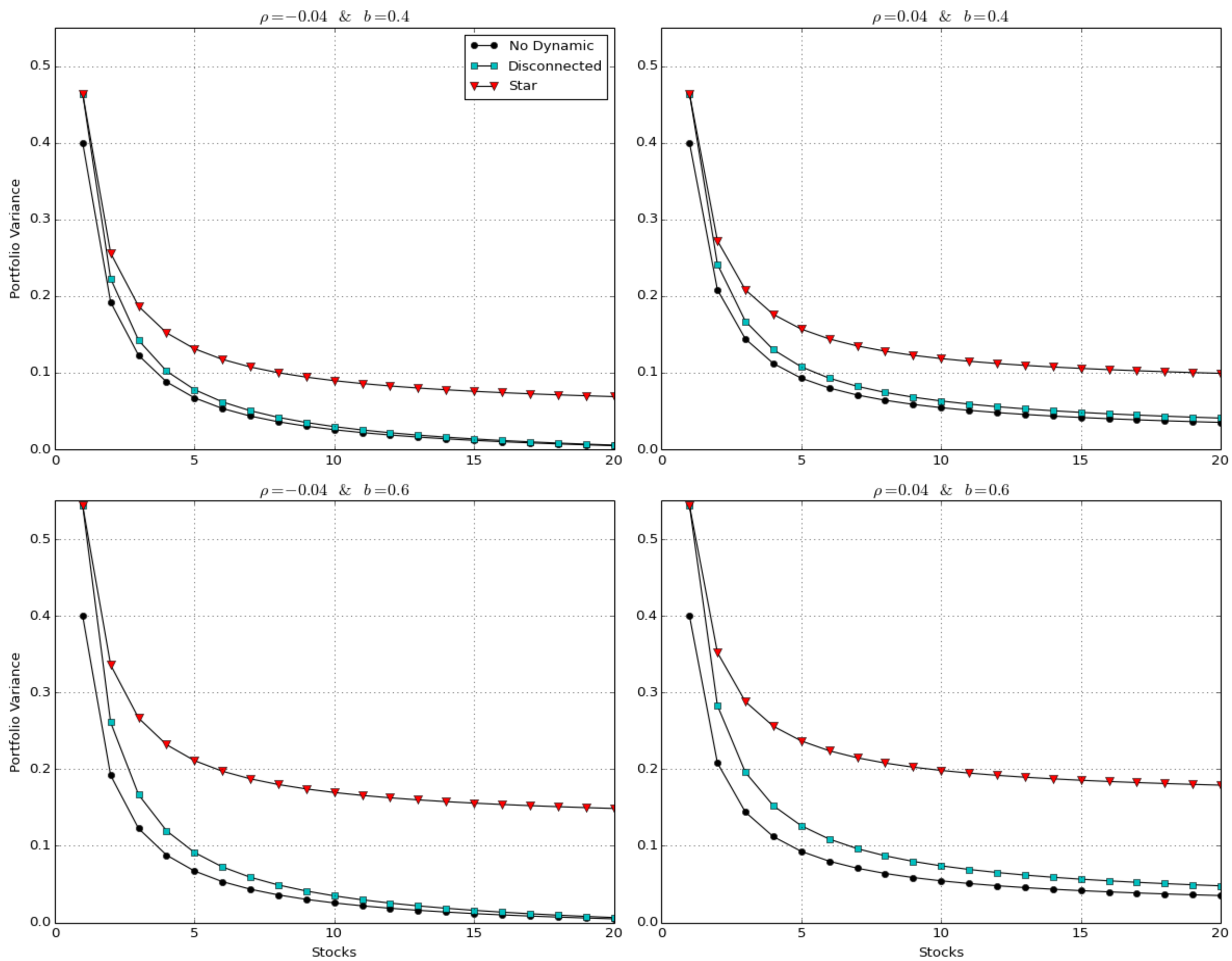

Figure 8. Portfolio variance for different topologies

Figure 8 plots four graphs depending on the values of the two fundamental parameters, specifically $\rho \in \{-0.04, 0.04\}$ and $b \in \{0.4, 0.6\}$. We consider positive values for b and it is expected for the negative values to lead to the same behavior of portfolio variance convergance. Additionally, since ρ is the mean correlation among n nodes, such parameter is not bounded from above but it is bounded from below since the minimum negative correlation coefficient among the

[9] Since portfolio variance of the circle network and inverse star network has the same expression as the disconnected network, their corresponding portfolio variance is not presented.

set of n variables is given by $-1/(n-1)$. Since the maximum portfolio size in figure 8 is 20, this explains the values given to ρ.[10]

There are three aspects that are worth to highlight. First, diversification benefits are evident given the negative slope shown by σ_p^2 in any of the network configurations and for any parameter specification. Thereby, larger portfolio size provides lower portfolio variance disregarding the network topology in place. Second, there is a clear ordering in terms of σ_p^2 for any given level of portfolio size that prevails irrespectively of the parameter specification. The worst performance is assigned to the star network and the best one corresponds to the case of no dynamics structure. Between these two extremes, the rest of the network configurations are located. Finally, we also observe the influence of quantity of b. From figure 8, it could be seen that for larger values of b, the portfolio variance of the disconnected network becomes closer to that of the star network and far away for the no-dynamical case.

In order to get further insights on the effects that stock network has on the diversification benefits, the portfolio variance elasticity $\xi(n)$ is defined as follows

$$\xi(n) = \frac{\partial \sigma_p^2(n)}{\partial n} \frac{n}{\sigma_p^2(n)} \tag{19}$$

Expression (20) and (21) provide the formulas for such elasticity for the disconnected network and star network, respectively.[11]

$$\xi^D(n) = \frac{\rho - 1}{\rho(n-1)+1} \tag{20}$$

$$\xi^S(n) = \frac{\rho - 1}{\rho(n-1)+b^2 n+1} \tag{21}$$

Figure 9 shows the behavior of diversification elasticity for the same specification of parameters as before.

10 Note that in the limiting case of $n \to \infty$, ρ should not be lower than zero.

11 The cases of no dynamics, circle and inverse star networks show the same elasticity given by equation (20).

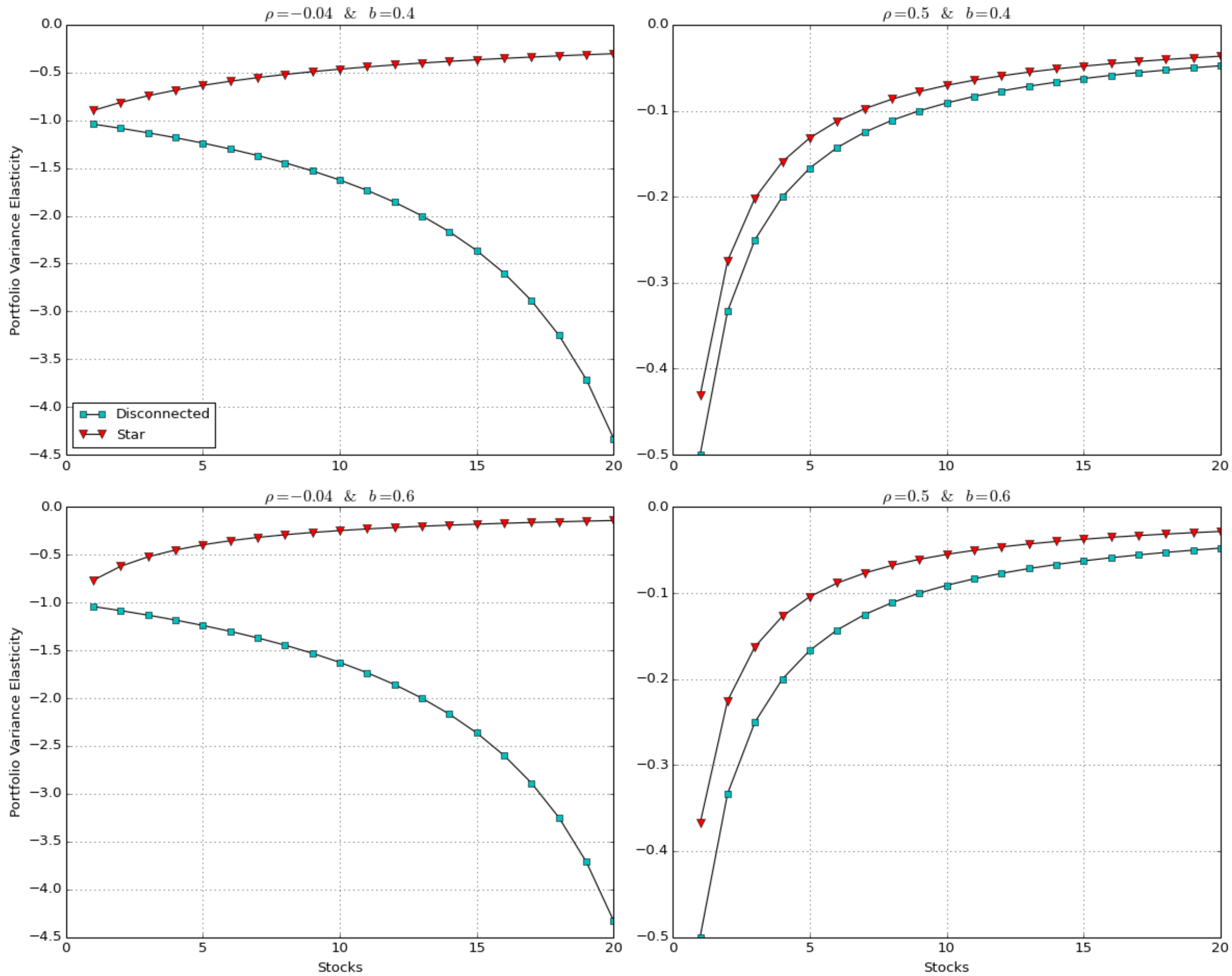

Figure 9. Variance elasticity for different topologies

Consistent with the figure 8, four aspects should be mentioned. First, ξ is negative stressing the benefit of diversification for any parameter specification and network configuration. Second, the ξ corresponding to star network is always lower (in absolute terms) than for the rest of the structures which relates to the existing potential of diversification embedded in different network architectures. Third, for $\rho > 0$, ξ shows a positive slope representing the decreasing marginal benefit of diversification. However, for $\rho < 0$, this behavior is preserved for the star network but this is not the case for the rest of the structures showing increasing benefit of diversification. Finally, higher b increases the difference between the benefit of diversification among the two types of structures.

As a summary it could be said that large concentration on the effects that a stock imposes to the rest of the system undermines the benefits of diversification, not only in term of its asymptotic limit but also for intermediate size portfolios. Additionally, this effect combined with negative mean correlation drastically changes the behavior of the portfolio variance for different values of n. Thereby, special attention should be put on the evolution of the network structures as a way to monitor the potential advantage of diversification.

5. Portfolio diversification and the impact of shocks in different structures

Until now, we have constructed a directed network for the stocks using a *VAR* model and accordingly, we discuss how various network topologies function with regard to portfolio diversification. In this section, we aim to investigate the impact of shocks on different stocks in various network topologies on portfolio return and accordingly on diversification benefits.

The portfolio return for period *t*, assuming weight vector $\boldsymbol{w}$, is computed as follows making use of the MA representation of a VAR process:

$$R_t^P = \boldsymbol{w}'\mu_r + \boldsymbol{w}'[\sum_{i=0}^{\infty} B^i u_{t-i}] \quad (22)$$

The portfolio return with respect to a one unit shock on stock *j*, *k* periods before is calculated as follows:

$$\frac{\partial R_t^P}{\partial u_{t-k,j}} = \sum_{i=1}^{N} w_i [B^k]_{ij} \quad (23)$$

Where $[B^k]_{ij}$ corresponds to the element in row i and column j in matrix B^k. The impact of shocks *k* periods before is transmitted through matrix B and as we increase *k*, we expect that stationarity assumption would drive this matrix to zero. Thereby, a shock on a stock far times before would have a slight impact on the portfolio return at time *t*. Moreover, the weights allocated to stocks is affected by stock *j* are also essential for determining the impact of the shock on the portfolio return.

In order to analyze the short-term influence of shocks on the diversification benefits, we introduce θ as the ratio between portfolio return and its standard deviation. We fixed the standard deviation to be the square root of long-term portfolio variance. Portfolio variance for different structures are computed in the preceding section. The impact of shocks to θ is driven by short-term portfolio returns. In this regard, the changes in θ with respect to a unit shock *k* periods before on stock *j* is computed as follows:

$$\frac{\partial\theta}{\partial u_{t-k,j}} = \frac{1}{\sigma_P} \times \frac{\partial R_t^P}{\partial u_{t-k,j}} \tag{24}$$

Where σ_P is the long-term portfolio variance. With regard to the effect of shocks in portfolio return through network topologies, we should analyze the elements of matrix B^k for increasing values of *k*.

For the star network, considering *SSA* to hold, the change in the value of θ in response to one unit shock to the central stock would be:

$$\frac{\partial\theta}{\partial u_{t-k,1}} = \frac{b^{2(k-1)}}{\sigma_P} \tag{25}$$

Since $|b| < 1$, as we increase *k*, the response to a shock decreases. Moreover, the impact of shocks on stocks in the peripheries on θ is zero since there is no path for this shock to be transmitted. The value computed in equation (25) is the change in θ as we shock the stock in the center by one unit in the star network. In case of negative shocks, the higher is the shock, the higher would be the reduction of the reduction in θ

In the case of disconnected network, the change in θ, under *SSA*, as we shock stock j is computed as follows:

$$\frac{\partial\theta}{\partial u_{t-k,1}} = \frac{1}{n}\frac{b^{2(k-1)}}{\sigma_P} \tag{26}$$

We observe that the impact of the shock to any stock on the ratio of return to risk is lower and as we increase the number of stocks in this type of network, the impact decreases.

For the full network, the change in θ, under *SSA*, is computed as follows:

$$\frac{\partial\theta}{\partial u_{t-k,1}} = \frac{3^{k-1}b^{2(k-1)}}{\sigma_P} \quad (27)$$

Since in the full network, all the stocks are connected to each other, we expect a shock to be transmitted rapidly and remains for a longer period. But it should be noted that the value of b for a full network should be low enough in order for the stationarity condition to hold. Additionally, as we will show in the empirical part, the higher is the number of outgoing and ingoing links, the lower would be the corresponding weights assigned to these links.

In order to demonstrate how the shocks impact the portfolio returns in short-term in different structures, we provide a numerical example. We consider three structures: disconnected network, star network and full network. Assuming *SSA*, we consider b to be equal to 0.4 in star and disconnected networks and equal to 0.04 in the full network structure. We consider ρ to be equal to 0.4 and variance to be 1 in all of the structures. We assume 100 stocks. The impact of a one unit shocks on stock *1* on portfolio return for 5 periods is presented in figure 10.

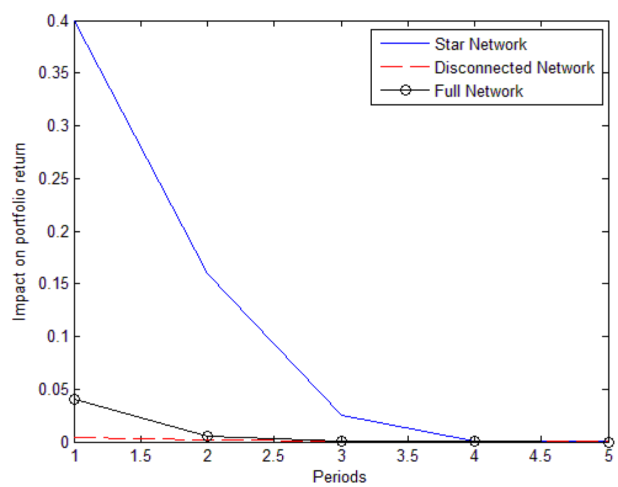

Figure 10. Change in theta for a one unit shock on a stock in the portfolio

The impact of a shock on the central stock in the star network last longer than that in full and disconnected network. Moreover, in the disconnected network, the shock does not have much impact on the portfolio return and with regard to full network, the impact of shocks disappear rapidly.

In total, we can argue that in the network structures where the distribution of out-going links is fat-talied, the impact of shocks on portfolio returns and consequently, on the portfolio diversification benefits is higher and long-lasting. In addition, to obtain a portfolio with lower short-term negative movement in relation to shocks in the stocks, it is advisable to construct the portfolio with a disconnected network structure.

5.1. Overall impact of shocks on portfolio performance

In order to analyze the overall impact of shocks, we employ the framework provided by Acemoglu et al. (2015). This framework helps us verify how the network interactions function in the event of shocks. Their methodology characterize the notion of systemic risk via network linkages in portfolio context.

In order to connect our model to the reduced form model of Acemoglu et al. (2015), we need to define (i) an interaction function, (ii) an interaction network and (iii) an aggregation function. In our context, we consider the interaction function to be a linear function of the form $f(x) = x$ and our interaction network defined by the directed network B. Moreover, the portfolio return would be our aggregation function. Thereby, the overall impact of a shock on stock i, on the return of the portfolio would be as follows:

$$\frac{\partial R_P}{\partial u_i} = w_i(I_N - B)^{-1}e_i$$

Where e_i is an *i-th* unit matrix with $N \times 1$ dimension. The element $(I_N - B)^{-1}$ corresponds to Bonacich centrality in network theory context. Bonacich centrality is higher for the stocks that either interact with a high number of stocks or with the stocks that are themselves highly central.

Accordingly, if the stocks are highly central, a shock to them would have higher impact on the return of the portfolio than low central ones. Moreover, the overall impact of shocks is also related to the amount of wealth we invest on the stocks. In total, using Bonacich centrality we can measure how systemic a stock is in a portfolio.

6. Empirical analysis

This section provides an empirical application of the concepts developed in the last sections. We rely on Datastream to account for adjusted returns of the 100 most capitalized constituents of the S&P-500 index. The period under analysis starts from 2004-11-02 until 2014-09-30. Since the goal is to characterize the evolution of the stock network, the dataset is divided in three sub-periods: Pre-Crisis from 2004-11-01 until 2008-02-29, Crisis from 2008-03-03 until 2009-03-31 and Post-Crisis from 2009-04-01 until 2014-09-30.

In order to estimate the VAR model in equation (1), we employ elastic net method of Zou and Hastie (2005). Tibshirani (1996) proposed LASSO algorithm that performs both parameter shrinkage and variable selection by including an l_1 penalty term. On the other hand, ridge regression performs similar to LASSO by minimizing sum of squared residuals subject to a penalty term. Ridge regression considers l_2 penalty term which exclude straightforward variable selection as in LASSO. The problem with LASSO is that it randomly selects a predictor from a group of correlated predictors. Elastic net combines both penalty terms in LASSO and ridge regression. In this way, it performs variable selection by shrinking parameters to zero and it is capable of distinguishing between predictors even when they are correlated[12].

In the first step, we apply elastic net to the returns in whole period and draw the histogram of weights in matrix *B* in order to see how the weights are distributed. The weights distribution is presented in figure 11. We observe that the weights are distributed around zero. This is expectable as elastic net tends to shrink the parameter estimation towards zero.

[12] Elastic net is explained in details in Appendix C

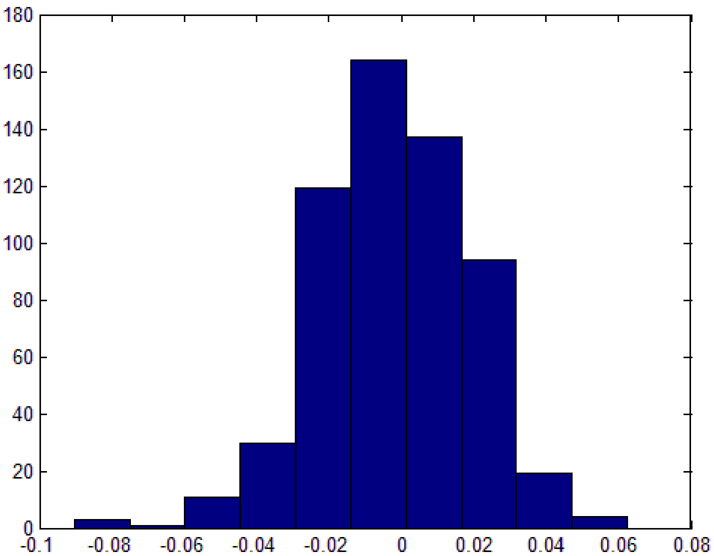

Figure 11. Weight distribution in serial dependence matrix *B*

In the next step, we investigate the relation between the number of connections a stock has in the network and the mean of the absolute values of the weights attached to links, stock strength. The results are presented in figure 12 where x-axis plots the connectivity of each stock and the y-axis the corresponding strength. A clear positive relationship is found for the whole period and for each of the subperiod under analysis except for the out degree case in the pre-crisis scenario.. Therefore, we can conclude that those stocks centrally placed in the network due to large number connectivity also shows larger individual effect upon the connected stocks

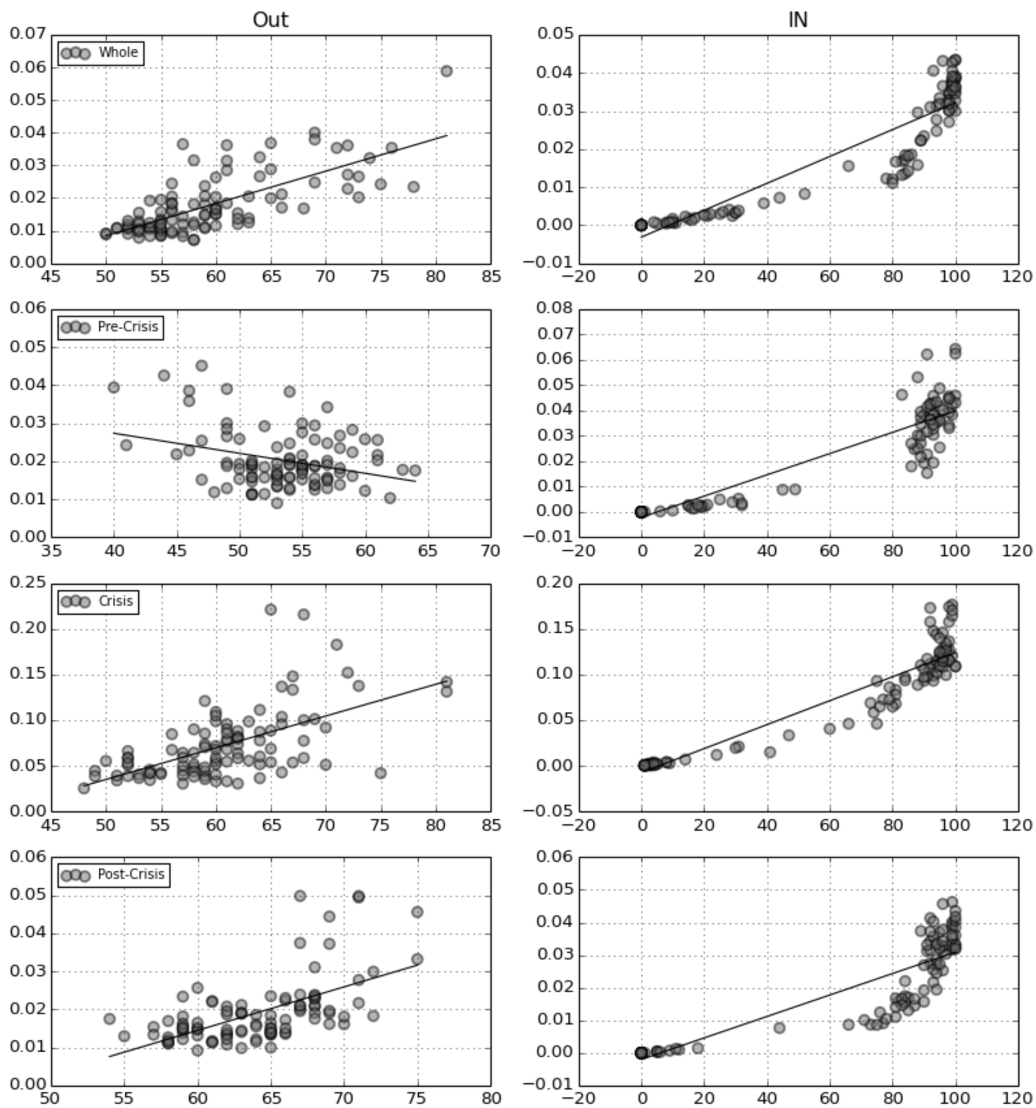

Figure 12. Number of in- and out-going edges and weight distribution

In the last sections, we observed a clear distinction between star network and the other network topologies. Thus, a fundamental aspect to measure is the extent to which the empirical network structure resembles a star-like the star network which could be done by mean of the concept of network centralization. The idea of network centralization is discussed in (Freeman 1978) and it is captured by C_x as follows:

$$C_x = \frac{\sum_{i=1}^{n}[C_x(p^*) - C_x(p_i)]}{\max \sum_{i=1}^{n}[C_x(p^*) - C_x(p_i)]} \tag{28}$$

where $C_x(p_i)$ is the centrality of node *i*, $C_x(p^*)$ is the largest node centrality in the network and $\max \sum_{i=1}^{n}[C_x(p^*) - C_x(p_i)]$ is the maximum possible sum of differences in node centrality for a network of size *n*. Naturally, in our case $C_x(p^*)$ is equal to the centrality of node *i* in the star network. Note that $0 < C_x < 1$ being 0 in case when every node is equally important and being 1 when the underlying network is the star network.

We still need to define precisely the centrality measures to be used in equation (28). Since our stock network is directed, it turns out that a particular node could be central in affecting or being affected. We rely on HITS algorithm (Kleinberg 1999) to quantify the threatening centrality and vulnerability centrality[13]. Note that from previous analysis, we now that the out-going links are relevant to diversification performance; thereby, the relevant centrality measure in the threatening centrality (see appendix D). Therefore the relevant centralization of the network in this case is as follows:

$$C_x = \frac{\sum_{i=1}^{n}[C_x(p^*) - C_x(p_i)]}{n-1} \tag{29}$$

We calculate this measure along with other measures of network characteristics for the datasets in different periods and the results are provided in table 1. We can verify the change in network structure during the crisis period. We see that the network becomes more connected as we go through the crisis. This is in concordance with the findings by Billio et al. (2012) where they show that the recent financial crisis lead to the network structure of stocks to become highly interconnected. Looking at the centralization measure, we see that during the crisis, the network is less like a star network in comparison to the network from before and after crisis periods.

[13] The full description of these centrality measures is presented in Appendix D.

	Networks thorugth time			
	Whole	Pre Crisis	Crisis	Post Crisis
Basics				
Nodes	100	100	100	100
Links	5958	5345	6069	6401
Density	0.60	0.54	0.61	0.65
Mean Degree	59.58	53.45	60.69	64.01
Distance				
Diameter	2	2	2	2
Mean Distance	1.16	1.21	1.15	1.13
Components				
Out	0.00	0.00	0.00	0.00
Strongly	0.84	0.72	1.00	0.79
In	0.16	0.28	0.00	0.21
Patterns of Connectivity				
Transitivity	0.87	0.84	0.88	0.90
Reciprocity	0.60	0.53	0.61	0.64
Assortativity	-0.28	-0.34	-0.28	-0.27
Centralization	0.32	0.28	0.24	0.29

* Measures corresponding to the unweithed version of the Stock Network
Table 1

In the following, we compute the threatening and vulnerability centralities at industrial levels presented in our datasets as it is described in appendix D. The results are shown in figure 13. Looking at the results for the whole dataset, it is interestingly to note that financial industry is highly threatening and additionally vulnerable. Moreover, we can also see that banks were highly vulnerable before the crisis and after the crisis, they become both highly threatening and vulnerable.

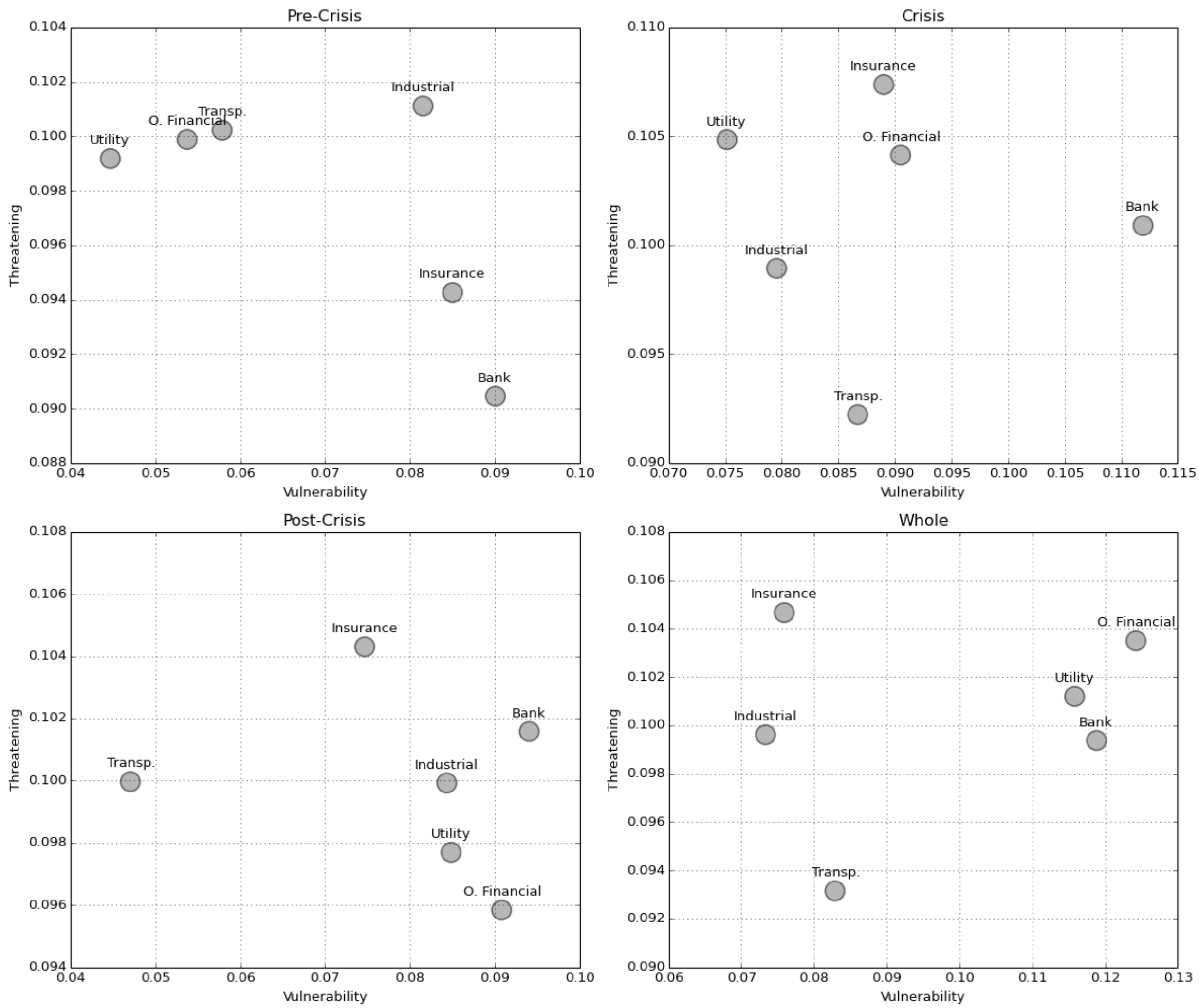

Figure 13. Threatening and vulnerability degrees for industries

7. Implication

We provide insights on how the topology structure of serial dependence of stocks would impact the portfolio variance and moreover, on how shocks would impact the portfolio return in short-term in different network topologies. Our findings have several implications. First, portfolio managers are able to build portfolios considering not only the variance-covariance matrix but also the network structure of stocks. In this way, they can increase the diversification benefits in their portfolios. Second, portfolio managers can build up portfolios that are less prone to sudden changes or shocks in the stocks by investing in portfolios toward disconnected or full network structure and away from star network structure.

Third, since we consider a general notion of portfolio for our analysis, the considered portfolios can represent a general stock exchange index where all of the stocks in that market are included in the portfolio. In this regard, we can analyze the impact of network structure on this market portfolio. Our calculations on the relation between network structure and portfolio variance can be extended to quantify the relationship between expected shortfall and network structure. We can compute the expected shortfall for the market assuming the market to be a portfolio of stocks. Expected shortfall has been considered as a measure of systemic risk (Acharya et al., 2010). Expected shortfall of the market is computed as the expected loss in the index conditional on this loss being greater than C (where C represents α level of portfolio return distribution). Subsequently, since the portfolio return follows a normal distribution, the expected shortfall for the portfolio would be as follows:

$$ES_t^{R_P}(\alpha) = w'(I_N - B)^{-1}a + \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha} \times (w'\Sigma_u w + w'B^1\Sigma_u B^{1'} w) \tag{30}$$

Where $\varphi(x)$ is the density of standard normal distribution. Accordingly, assuming no constants in the VAR model, we can conclude that there is a direct relationship between expected shortfall of the index and variance of the portfolio.

$$ES_t^{R_P}(\alpha) \sim (w'\Sigma_u w + w'B\Sigma_u B'w) \tag{31}$$

Thereby, the higher the variance, the higher would be the expected shortfall of the market index. We can conclude that the star network structure have a higher variance level and thereby, a higher expected shortfall level and additionally, in the short-run, the impact of a shock is both higher and preservative than other structures. In this regard, a stock exchange that depicts a star network structure caries higher systemic risk than a full network or an individual network structure.

8. Conclusion

In this paper, we construct a directed network of stocks by estimating the one period serial dependency matrix via a *VAR(1)* model. We compare the portfolio variance across several special network topologies for a naïve diversification strategy and we find that larger asymmetry in the number of out-going links across stocks leads to lower diversification benefits. We find that for individual, circle and full network structures, the portfolio variance acts in the same manner asymptotically while in the star network, the portfolio variance does not capture the minimum possible values in other network structures.

Moreover, we investigate the impact of shocks on the short-term diversification benefits and we find that for the star network structure, the impact of a shock on the central stock is higher and longer-lasting.

In the empirical part, we analyze the portfolio returns during the 2007/2008 crisis period. We find that after crisis, the network structure is digressing toward a star network structure and during the crisis, the stock network structure is denser. Moreover, we also find that the higher is the number of connections for a stock, the larger would be the mean absolute value of the weights attached to the links

For the future research, we propose analyzing numerically the impact of the directed network structure on the stability of the market. Considering various special cases as we include in our paper, we can verify how the market, considered as a portfolio index, would react to shocks in different parts of the network.

References

Acemoglu, D. et al., 2012. The Network Origins of Aggregate Fluctuations. *Econometrica*, 80(5), pp.1977–2016.

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Networks, Shocks, and Systemic Risk. No. w20931. National Bureau of Economic Research, 2015.

Allen, Franklin and Douglas Gale (2000), "Financial contagion." Journal of Political Economy, 108, 1–33.

Barberis, N. 2000. Investing for the long run when returns are predictable. Journal of Finance 55:225–64

Barigozzi, M. & Brownlees, C., 2013. Nets: Network Estimation for Time Series. *SSRN - Working Paper*, pp.1–43.

Billio, M. et al., 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104(3), pp.535–559.

Bird, R. & Tippett, M., 1986. Note — Naive Diversification and Portfolio Risk — A Note. *Management Science*, 32(2), pp.244–251.

Bloomfield, Ted, Richard Leftwich, and John B. Long Jr. "Portfolio strategies and performance." Journal of Financial Economics 5.2 (1977): 201-218.

Bonanno, G. et al., 2004. Networks of equities in financial markets. *The European Physical Journal B - Condensed Matter*, 38(2), pp.363–371.

Bonanno, G., Lillo, F. & Mantegna, R.N., 2001. High-frequency cross-correlation in a set of stocks. *Quantitative Finance*, 1(1), pp.96–104.

Campbell, J. et al., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance*, 56(1), pp.1–43.

Campbell, J.Y.,Y. L. Chan, and L. M. Viceira. 2003.Amultivariate model of strategic asset allocation. Journal of Financial Economics 67:41–80.

Campbell, J. Y., and L. M. Viceira. 1999. Consumption and portfolio decisions when expected returns are time varying. Quarterly Journal of Economics 114:433–95.

Campbell, J. Y., and L. M. Viceira. 2002. Strategic asset allocation. NewYork: Oxford University Press.

Chen, S. & Keown, A., 1981. Risk Decomposition and Portfolio Diversification When Beta is Nonstationary: A Note. *The Journal of Finance*, 36(4), pp.941–947.

Chordia, T., and B. Swaminathan. 2000. Trading volume and cross-autocorrelations in stock returns. Journal of Finance 55:913–35.

DeMiguel, V., Nogales, F.J. & Uppal, R., 2014. Stock Return Serial Dependence and Out-of-Sample Portfolio Performance. *Review of Financial Studies*, 27(4), pp.1031–1073. Available at: http://rfs.oxfordjournals.org/cgi/doi/10.1093/rfs/hhu002 [Accessed October 6, 2014].

Diebold, F.X. & Yılmaz, K., 2014. On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1), pp.119–134.

Domian, D., Louton, D. & Racine, M., 2007. Diversification in portfolios of individual stocks: 100 stocks are not enough. *Financial Review*, 42(4), pp.557–570.

Elton, E. & Gruber, M., 1977. Risk Reduction and Portfolio Size : An Analytical Solutiont. *The Journal of Business*, 50(4), pp.415–437.

Evans, J. & Archer, S., 1968. Diversification and the Reduction of Dispersion: An Empirical Analysis. *The Journal of Finance*, 23(5), pp.761–767.

Eun, C. S., and S. Shim. 1989. International transmission of stock market movements. Journal of Financial and Quantitative Analysis 24:241–56.

Freeman, L., 1978. Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1(3), pp.215–239.

Freixas, Xavier, Bruno M. Parigi, and Jean-Charles Rochet (2000), "Systemic risk, interbank relations, and liquidity provision by the central bank." Journal ofMoney, Credit and Banking, 32, 611–638.

Garas, A. & Argyrakis, P., 2007. Correlation study of the Athens Stock Exchange. *Physica A: Statistical Mechanics and its Applications*, 380, pp.399–410.

Hautsch, N., Schaumburg, J. & Schienle, M., 2014a. Financial Network Systemic Risk Contributions. *Review of Finance*.

Hautsch, N., Schaumburg, J. & Schienle, M., 2014b. Forecasting systemic impact in financial networks. *International Journal of Forecasting*, 30(3), pp.781–794.

Huang, W.Q., Zhuang, X.T. & Yao, S., 2009. A network analysis of the Chinese stock market. *Physica A: Statistical Mechanics and its Applications*, 388(14), pp.2956–2964.

Jackson, M.O., 2010. *Social and Economic Networks*, Princeton University Press.

James, J., Kasikov, K. & Edwards, K.-A., 2012. The end of diversification. *Quantitative Finance*, 12(11), pp.1629–1636.

Jung, W. et al., 2006. Characteristics of the Korean stock market correlations. *Physica A: Statistical Mechanics and its Applications*, 361(1), pp.263–271.

Kleinberg, J., 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5), pp.604–632.

Klemkosky, R. & Martin, J., 1975. The Effect of Market Risk on Portfolio Diversification. *The Journal of Finance*, 30(1), pp.147–154.

Lütkepohl, H., 2007. *New Introduction to Multiple Time Series Analysis*, Springer. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:New+introduction+to+multiple+time+series+analysis#0 [Accessed October 19, 2014].

Mantegna, R.N., 1999. Hierarchical structure in financial markets. *The European Physical Journal B - Condensed Matter*, 11(1), pp.193–197.

Mao, J., 1970. Essentials of Portfolio Diversification Strategy. *The Journal of Finance*, 25(5), pp.1109–1121.

Marín, J.M. & Rubio, G., 2011. *Economia Financiera*, Antoni Bosch.

Markowitz, H., 1952. Portfolio Selection. *The Journal of Finance*, 7(1), pp.77–91.

Newman, M.E.J., 2010. *Networks: An Introduction*, Oxford University Press.

Onnela, J.P. et al., 2003. Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, 68(5), p.056110.

Ozsoylev, H., Walden, J. & Yavuz, M.D., 2014. Investor Networks in the Stock Market. *Review of Financial Studies*, 27(5), pp.1323–1366.

Pozzi, F., Di Matteo, T. & Aste, T., 2013. Spread of risk across financial markets: better to invest in the peripheries. *Scientific reports*, 3, p.1665.

Rapach, D., Strauss, J. & Zhou, G., 2013. International Stock Return Predictability: What Is the Role of the United States? *The Journal of Finance*, 68(4), pp.1633–1662. Available at: http://doi.wiley.com/10.1111/jofi.12041 [Accessed March 24, 2014].

Samuelson, P., 1967. General proof that diversification pays. *Journal of Financial and Quantitative Analysis*, 2(1), pp.1–13.

Statman, M., 1987. How many stocks make a diversified portfolio? *Journal of Financial and Quantitative Analysis*, 22(3), pp.353–363.

Tumminello, M., Lillo, F. & Mantegna, R.N., 2010. Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior & Organization*, 75(1), pp.40–58.

Vandewalle, N., Brisbois, F. & Tordoir, X., 2001. Non-random topology of stock markets. *Quantitative Finance*, 1(3), pp.372–374.

Zou, H. & Hastie, T., 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), pp.301–320.

Zou, H. & Zhang, H.H., 2009. On the Adaptive Elastic-Net With a Diverging Number of Parameters. *Annals of statistics*, 37(4), pp.1733–1751.

Appendix A: Convergance of serial dependence matrix under stationarity condition

In order to demonstrate how rapidly the powers of matrix B converge to zero, we consider 10 highest capitalized stocks and calculate the matrix B. Next, we compute the sum of absolute values of elements for different powers of matrix B up to the tenth power. The results are presented in the following figure:

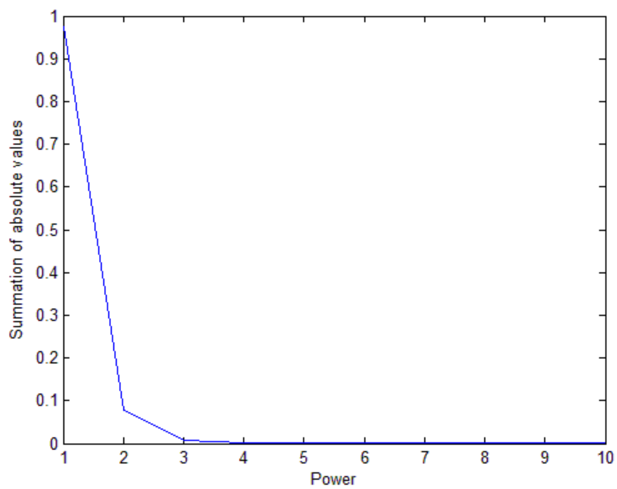

Figure A1: Summation of absolute values for different powers of matrix B

Appendix B: Influence of each motif on portfolio variance

In case (a), a stock is affecting by itself. This contributes to portfolio variance with the following expression: $(w_i)^2 {B_{ii}}^2 {\sigma_i}^2$ where ${\sigma_i}^2$ is the variance of stock i. In case (b) in figure 2, two different stocks affect two separate stocks contributing to the portfolio variances as follows:

$$w_k w_l B_{ki} B_{lj} \sigma_{ij} + (w_k)^2 B_{ki} B_{ki} {\sigma_i}^2 + (w_l)^2 B_{lj} B_{lj} {\sigma_j}^2 \qquad \text{(A1)}$$

Intuitively, the weights invested in stocks with in-going links and also the covariance between out-going stocks and their individual variance are the determinant of influence on portfolio variance.

In case (c), stock i is affecting both stocks l and k. The impact of this interaction on portfolios variance is

$$w_k w_l B_{ki} B_{li} {\sigma_i}^2 + (w_k)^2 B_{ki} B_{ki} {\sigma_i}^2 + (w_l)^2 B_{li} B_{li} {\sigma_i}^2 \qquad \text{(A2)}$$

Under this pattern of connectivity, the variance of the initiator stock plays the important role. This is straightforward as we can see that any perturbation in the prices of both stocks k and l comes from changes in return of the stock that is threatening them.

In case (d), both stocks are impacting each other. The influence of this dynamic structure is:

$$w_i w_k B_{ki} B_{ik} \sigma_{ik} + (w_k)^2 B_{ki} B_{ki} {\sigma_i}^2 + (w_i)^2 B_{ik} B_{ik} {\sigma_k}^2 \qquad \text{(A3)}$$

The individual variance of both stocks and also their covariance is determinant in quantifying the portfolio variance.

In case (e), two stocks i and j are affecting one stock k. The notion signifying this interaction in portfolio's variance is as follows:

$$(w_k)^2 B_{ki} B_{kj} \sigma_{ij} + (w_k)^2 B_{ki} B_{ki} {\sigma_i}^2 + (w_k)^2 B_{kj} B_{kj} {\sigma_j}^2 \qquad \text{(A4)}$$

The weight allocated to stock *k* and also the variance and covariance of out-going stocks are the major players in this motif.

Finally, in case (f), the underlying portfolio variance impact would be:

$$(w_k)^2 B_{ki} B_{ki} {\sigma_i}^2 + (w_j)^2 B_{jk} B_{jk} {\sigma_k}^2 \qquad \text{(A5)}$$

Appendix C: Elastic net estimation procedure

The equation for stock *i* in the *VAR(1)* model is represented as follows:

$$r_{t+1}^i = a^i + \sum_{j=1}^N B_{ij} r_t^j + u_{t+1}^i \qquad \text{(A6)}$$

The values for row *i* of matrix *B* is estimated as follows:

$$\min_{B_{ij}} \left[\sum_{t=1}^{T} (r_{t+1}^{i} - \sum_{j=1}^{N} B_{ij} r_t^j)^2 + \lambda_1 \sum_{j=1}^{N} |B_{ij}| + \lambda_2 \sum_{j=1}^{N} B_{ij}^2 \right] \tag{A7}$$

By estimating the above expression, we compute the row *i* of matrix *B*. To find out all the elements of matrix *B*, we estimatate via elastic net for each row by changing the dependent stocks. The regularization parameters λ_1 and λ_2 refer to the *l1* and *l2* penalty terms, respectively. These two values are estimated using 10-fold cross validation.

Appendix D: Threatening and vulnerability centrality

Following (Kleinberg 1999) HITS algorithm is applied. In a directed network the direction of the links is relevant. Thus, each node has in principle to roles that can be captured by the centrality of its position in such roles, vulnerability centrality, v_i, and threatening centrality t_i. An stock becomes a high threat to the system as long as it *points to* many stock that are highly vulnerable. On the other hand, a stock is highly vulnerable as long as it is *pointed to* by many threatening stocks. In this approach, the vulnerability centrality is proportional to the sum of the threatening centralities of the nodes that point to it

$$v_i = \alpha_v \sum_j e_{ij} t_j \tag{A8}$$

where α_v is a constant. Which in matrix notation is given by

$$v = \alpha_v E t \tag{A9}$$

Similarly, the threatening centrality is proportional to the sum of the vulnerability centrality of the nodes it points to

$$t_i = \alpha_t \sum_j e_{ji} v_j \tag{A10}$$

Where α_t is another constant. In matrix notation

$$t = \alpha_t E^T v \tag{A11}$$

Combining (25) and (27) we get

$$v = \alpha E E^T v \tag{A12}$$

$$t = \alpha E^T E t \tag{A13}$$

Where $\alpha = \alpha_v \alpha_t$. Thus, the vulnerability centrality and the threatening centrality are respectively given by eigenvectors of EE^T and $E^T E$ corresponding to the largest eigenvalues. Note the such matrices have the same eigenvalues. Note that each stocks has 2 characteristics, vulnerable and threatening.

For the case of star network, it could be proved that threatening centrality for the center is one and zero for the rest of the stocks. However, the vulnerability centrality is the same value for each element (including the center) when the network has loop or is the same number for all the element expect the center that assumes zero vulnerability when there is no loop