Journal of Empirical Finance 38 (2016) 157–180

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

A network approach to portfolio selection

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

[a] Department of Research and Statistics at CNMV, Calle Edison, 4, 28006 Madrid, Spain
[b] Department of Business Administration at Universidad Carlos III de Madrid, Calle Madrid, 126, 28903, Getafe, Madrid, Spain

ARTICLE INFO

Article history:
Received 28 May 2015
Received in revised form 27 May 2016
Accepted 3 June 2016
Available online 07 June 2016

Keywords:
Network theory
Centrality
Portfolio selection

ABSTRACT

In this study, a financial market is conceived as a network where the securities are nodes and the links account for returns' correlations. We theoretically prove the negative relationship between the centrality of assets in this financial market network and their optimal weights under the Markowitz framework. Therefore, optimal portfolios overweight low-central securities to avoid the large variances that result when highly influential stocks are included in the investor's opportunity set. Next, we empirically investigate the major financial and market determinants of stock's centralities. The evidence indicates that highly central nodes tend to coincide with older, larger-cap, cheaper and financially riskier securities. Finally, we explore by means of in-sample and out-of-sample analysis the extent to which the structure of the stock market network can be employed to improve the portfolio selection process. We propose a network-based investment strategy that outperforms well-known benchmarks while presenting positive and significant Carhart alphas. The major contribution of the paper is to employ the financial market network as a useful device to improve the portfolio selection process by targeting a group of assets according to their centrality.

1. Introduction

In his seminal paper, Markowitz (1952) laid the foundation of modern portfolio theory. In this static framework, investors optimally allocate their wealth across a set of assets considering only the first and second moment of the returns' distribution. Despite the profound changes derived from this publication, the out-of-sample performance of Markowitz's prescriptions is not as promising as expected. The poor performance of Markowitz's rule stems from the large estimation errors on the vector of expected returns (Merton, 1980) and on the covariance matrices (Jobson and Korkie, 1980) leading to the well-documented error-maximizing property discussed by Michaud and Michaud (2008). The magnitude of this problem is evident when we acknowledge the modest improvements achieved by those models specifically designed to tackle the estimation risk (DeMiguel et al., 2009). Moreover, the evidence indicates that the simple yet effective equally-weighted portfolio rule has not been consistently out-performed by more sophisticated alternatives (Bloomfield et al., 1977; DeMiguel et al., 2009; Jorion, 1991).

Recently, researchers from different fields have characterized financial markets as networks in which securities correspond to the nodes and the links relate to the correlation of returns (Barigozzi and Brownlees, 2014; Billio et al., 2012; Bonanno et al., 2004; Diebold and Yilmaz, 2014; Hautsch et al., 2015; Mantegna, 1999; Onnela et al., 2003; Peralta, 2015; Tse et al., 2010; Vandewalle et al., 2001; Zareei, 2015). In spite of the novel and interesting insights obtained from these network-related papers, most of their results are fundamentally descriptive and lack concrete applications in portfolio selection process. We contribute to

* Corresponding author.
E-mail addresses: gperalta@cnmv.es (G. Peralta), azareei@emp.uc3m.es (A. Zareei).

http://dx.doi.org/10.1016/j.jempfin.2016.06.003
0927-5398/

this line of research by investigating the extent to which the underlying structure of this financial market network can be used as an effective tool in enhancing the portfolio selection process.

Our theoretical results establish a bridge between Markowitz's framework and the network theory. On the one hand, we show a negative relationship between optimal portfolio weights and the *centrality* of assets in the financial market network. The intuition is straightforward: those securities that are strongly embedded in a correlation-based network greatly affect the market and their inclusion in a portfolio undermines the benefit of diversification resulting in larger variances. We refer to the centrality of stocks as their systemic dimension. On the other hand, each security is also characterized by an individual dimension such as Sharpe ratio or volatility depending on the specific portfolio formation objective. Next, we theoretically show a positive relationship between the assets' individual performances and their optimal portfolio weights. In a nutshell, optimal weights from the Markowitz framework can be interpreted as an optimal trade-off between the securities' systemic and individual dimensions in which the former is intimately related to the notion of network centrality.

From a descriptive perspective and relying on US data, we present evidence indicating that financial stocks are the most central nodes in the financial market network in accordance with Peralta (2015) and Tse et al. (2010). Additionally, we document a positive association between the centrality of a security and its corresponding beta from CAPM pointing out the large, although not perfect, correlation between this network indicator and the standard measure of systematic risk. In order to identify the salient financial and market features affecting securities' centrality, we estimate several specifications of a quarterly-based panel regression model upon a set of 200 highly capitalized stocks in the S&P500 index from Oct-2002 to Dec-2012. Our results present some empirical evidence indicating that highly central stocks correspond to old firms with low prices, great market capitalizations and low cash holdings.

Finally, by means of in-sample and out-of-sample analysis, we investigate the extent to which the structure of the financial market network can be used to enhance the portfolio selection process. In order to check the robustness of our results and to avoid data mining bias, four datasets are considered, accounting for different time periods and markets. We propose a network-based investment rule, termed as ρ-dependent strategy, and report its performance against well-known benchmarks. The evidence shows that our network-based strategy provides significant larger out-of-sample Sharpe ratios compared to the ones obtained by implementing the $1/N$ rule or Markowitz-based models. Moreover, these enhanced out-of-sample performances are not explained by large exposures to the standard risk factors given the reported positive and statistically significant Carhart alphas. Finally, it is worth mentioning that our results are robust to different portfolio settings and transaction cost. We argue that our network-based investment policy captures the *logic* behind Markowitz's rule while making more efficient use of fundamental information, resulting in a substantial reduction of wealth misallocation.

The contribution of the paper is twofold. On the one hand, this paper sheds light on the connection between the modern theory of portfolios and the emerging literature on financial networks. On the other hand, our network-based investment strategy attempts to simplify the portfolio selection process by targeting a group of stocks within a certain range of network centrality. As far as we are aware, Pozzi et al. (2013) is the only paper that attempts to take advantage of the topology of the financial market network for investment purposes. They argue in favor of an unconditional allocation of wealth towards the outskirts of the structure. We depart from their results by proposing the ρ-dependent strategy which is contingent on the correlation between the systemic and individual dimensions of the assets comprising the financial market network. Moreover, and in contrast to their synthetic centrality index, we show that our measure of centrality is strongly rooted in the principles of portfolio theory.

The remainder of the paper is organized as follows. Section 2 presents the notion of assets' centrality in the financial network and its connection to Markowitz framework. Section 3 describes the datasets used in the empirical applications. Section 4 provides a detailed statistical description of stocks in accordance to their centrality. Section 5 addresses the interaction between assets' centrality and optimal portfolio weights by relying on an in-sample analysis. Section 6 presents the ρ-dependent strategy and compares its out-of-sample performance to various conventional portfolio strategies. Finally, Section 7 concludes and outlines future research lines.

2. A bridge between optimal portfolio weights and network centrality

The notion of centrality, intimately related to the social network analysis, aims to quantify the influence/importance of certain nodes in a given network. As discussed in Freeman (1978), there are several measurements in the literature each corresponding to the specific definition of centrality. The so-called *eigenvector centrality*, firstly proposed by Bonacich (1972), has become standard in network analysis. This section formally defines this measure and determines its relationship with optimal portfolio weights.

2.1. Defining network centrality

We denote by $G=\{N,\omega\}$, a network composed by a set of nodes $N=\{1,2,\ldots,n\}$ and a set of links, ω, connecting pairs of nodes. If there is a link between nodes i and j, we indicate it as $(i,j)\in\omega$. A convenient rearrangement of the network information is provided by the $n\times n$ adjacency matrix $\Omega=[\Omega_{ij}]$ whose element $\Omega_{\mathrm{ij}}\neq 0$ whenever $(i,j)\in\omega$. The network G is said to be undirected if no-causal relationships are attached to the links implying that $\Omega=\Omega^{\mathrm{T}}$ since $(i,j)\in\omega\Longleftrightarrow(j,i)\in\omega$. When Ω_{ij} entails a causal association from node j to node i, the network G is said to be directed. In this case, it is likely that $\Omega\neq\Omega^{T}$ since $(i,j)\in\omega$ does not necessarily imply $(j,i)\in\omega$. For unweighted networks, $\Omega_{ij}\in\{0,1\}$ and therefore only on/off relationships exist. On the contrary, when

$\Omega_{ij}\mathbb{R}$, the links track the intensity of the interactions between nodes giving rise to weighted networks. The reader is referred to Jackson (2010) for a comprehensive treatment of the network literature.

According to Bonacich (1987, 1972), the eigenvector centrality of node i, denoted by v_i, is defined as the proportional sum of its neighbors' centrality.[1] (Newman, 2004) extends this notion to weighted networks for which v_i is proportional to the weighted sum of the centralities of neighbors of node i with Ω_{ij} as the corresponding weighting factors. It is computed as follows.

$$v_i \equiv \lambda^{-1}\sum\nolimits_j \Omega_{ij} v_j \tag{1}$$

Note that node i becomes highly central (large v_i) by being connected either to many other nodes or to just few highly central ones. By restating Eq. (1) in matrix terms, we obtain $\lambda v = \Omega v$ indicating that the centrality vector v is given by the eigenvector of Ω corresponding to the largest eigenvalue, λ.[2] More formally:

Definition 1. Consider the undirected and weighted network $G = \{N, \Omega\}$, with N as the set of nodes and Ω as the adjacency matrix. The eigenvector centrality of node i in G denoted by v_i is proportional to the i-th component of the eigenvector of Ω corresponding to the largest eigenvalue λ_1.

2.2. Key results from the modern portfolio theory

The mathematical principles of modern portfolio theory were established in (Markowitz, 1952). Given that our theoretical results strongly rely on this framework, this section briefly reviews two of its fundamental results: the minimum-variance and the mean-variance investment rules.

Let us assume n risky securities with expected returns vector, μ, and covariance matrix, $\Sigma = [\sigma_{ii}]$. Consider the problem of finding the vector of optimal portfolio weights, w, that minimizes the portfolio variance subject to $w^T \mathbf{1} = 1$ where 1 (in bold) corresponds to a column vector whose components are equal to one. This strategy is commonly known as minimum-variance or *minv* for short. Formally the problem is stated as:

$$min_w \sigma_p^2 = w^T \Sigma w \quad subject\ to \quad w^T 1 = 1 \tag{2}$$

The solution of Eq. (2) is given by:

$$w^*_{minv} = \frac{1}{\mathbf{1}^T \Sigma^{-1} 1} \Sigma^{-1} 1 \tag{3}$$

Denoting by Ω the correlation matrix of returns, and by Δ the diagonal matrix whose ith-main diagonal element is $\sigma_i = \sqrt{\sigma_{ii}}$, the relationship between Ω and Σ can be written as $\Sigma = \Delta\Omega\Delta$. Then, Eq. (3) is restated in terms of the correlation matrix as follows.

$$\hat{w}^*_{minv} = \varphi_{minv} \Omega^{-1} \epsilon \tag{4}$$

where $\hat{w}^*_{i,minv} = w^*_{i,minv} * \sigma_i$, $\varphi_{minv} = \frac{1}{1^T \Sigma^{-1} 1}$ and $\epsilon_i = 1/\sigma_i$.

The introduction of a risk-free security whose return is given by r_f allows us to account for the mean-variance investment rules. We denote the excess return of security i by $r_i^e \equiv r_i - r_f$ and the vector of expected excess returns by μ^e. The problem of finding the optimal portfolio weights that minimize the portfolio variance for a given level of the portfolio expected excess return R^e is established as follows:

$$\min_w \sigma_p^2 = w^T \Sigma w \ subject\ to\ w^T \mu^e = R^e \tag{5}$$

The strategy implied by Eq. (5) is commonly known in the financial literature as the mean-variance strategy or *mv* for short. Since an investor's wealth might be partially allocated to the risk-free security and short sales of the risk-free security are allowed, the restriction $w^T 1 = 1$ is not included in Eq. (5).[3] The optimum mean-variance portfolio weights vector is computed as follows:

$$w^*_{mv} = \frac{R^e}{\boldsymbol{\mu}^{e\boldsymbol{T}} \Sigma^{-1} \mu^e} \Sigma^{-1} \mu^e \tag{6}$$

[1] The terms eigenvector centrality and centrality are used interchangeably throughout this study.

[2] In principle, each eigenvector of Ω is a solution to Eq. (1). However, the centrality vector corresponding to the largest component in the network is given by the eigenvector corresponding to the largest eigenvalue (Bonacich, 1972).

[3] Nevertheless, when the tangency portfolio is considered, $w^T 1 = 1$ must hold.

Following the same reasoning as before, Eq. (6) is written in terms of the correlation matrix as follows:

$$\hat{w}^*_{mv} = \varphi_{mv}\Omega^{-1}\hat{\mu}^e \tag{7}$$

where $\hat{w}^*_{i,mv} = w^*_{i,mv} * \sigma_i$, $\varphi_{mv} = \frac{R^e}{\mu^{eT}\Sigma^{-1}\mu^e}$ *and* $\hat{\mu}^e_i = \mu^e_i/\sigma_i$.

2.3. The relationship between optimal portfolio weights and asset centralities

By interpreting the correlation matrix of returns as the adjacency matrix of a given network, an overlapping region between portfolio theory and network theory is established. More formally:

Definition 2. Consider N to be a set of securities in a given asset opportunity set and Ω the corresponding returns' correlation matrix. The undirected and weighted financial market network is $FMN = \{N, \Omega\}$, with N as the set of nodes and Ω as the adjacency matrix.

Throughout the study, we set the main diagonal of Ω to zero in order to discard meaningless self-loops in a given financial market network. Since the eigenvectors' structures and the ordering of eigenvalues are the same after performing this operation, our statements in terms of eigenvector centrality remain valid (see appendix A).

Proposition 1 and Corollary 1 establish the negative relationship between security i's optimal portfolio weight and its centrality in the respective financial market networks for both the *minv* and *mv* strategies. The reader is referred to appendix A for a detailed proof.

Proposition 1. Consider a financial market network $FMN = \{N, \Omega\}$ where $\{v_1, \ldots v_n\}$ and $\{\lambda_1, \ldots \lambda_n\}$ account for the sets of eigenvectors and eigenvalues (in descending order) of Ω, respectively. The optimal portfolio weights in the Eqs. (4) and (7) can be written as:

$$\hat{w}^*_{minv} = \varphi_{minv}\epsilon + \varphi_{minv}\left(\frac{1}{\lambda_1} - 1\right)\epsilon_M v_1 + \Gamma_{minv} \tag{8}$$

$$\hat{w}^*_{mv} = \varphi_{mv}\hat{\mu}^e + \varphi_{mv}\left(\frac{1}{\lambda_1} - 1\right)\hat{\mu}^e_M v_1 + \Gamma_{mv} \tag{9}$$

where $\epsilon_M = ({v_1}^T\epsilon)$, $\Gamma_{minv} = \varphi_{minv}[\sum_{k=2}^{n}(\frac{1}{\lambda_k} - 1)v_k v_k^T]\epsilon$, $\hat{\mu}^e_M = {v_1}^T\hat{\mu}^e$ *and* $\Gamma_{mv} = \varphi_{mv}[\sum_{k=2}^{n}(\frac{1}{\lambda_k} - 1)v_k v_k^T]\hat{\mu}^e$.

Note that ϵ_M and $\hat{\mu}^e_M$ in Eqs. (8) and (9) account for weighted averages of the inverted standard deviations of returns and Sharpe ratios, respectively, with weighting factors given by the elements of v_1. From a principal component perspective, we interpret them as the corresponding variables at market level. Moreover, since the empirical evidence indicates that only eigenvector elements corresponding to the largest eigenvalue have informational content (Green and Hollifield, 1992; Laloux et al., 1999; Trzcinka, 1986), we focus on v_1 and λ_1 by defining Γ_{minv} and Γ_{mv} in terms of v_j and λ_j for $j>1$.

The first term in Eqs. (8) and (9) considers simple investment rules that only take into account the performance of securities as if they were in isolation. Therefore, lower (higher) standard deviations of returns (Sharpe ratios) are consistent with higher optimal portfolio weights. We call this the *individual dimension of securities*. The second term in the same expressions quantifies the extent to which optimal weights deviate from the previously mentioned rule due to the centrality of securities. We call this the *systemic dimension of securities*. Corollary 1 states that, under plausible conditions, there is a negative relationship between optimal portfolio weights and network centralities.

Corollary 1. Assuming that $\lambda_1 > 1$ and $\epsilon_M, \hat{\mu}^e_M \in \mathbb{R}_+$. Then, $\frac{\partial \hat{w}^*_{r,i}}{\partial v_{1,i}} < 0$ for $r = minv, mv$.

Based on Rayleigh's inequality (Van Mieghem, 2011), the assumption $\lambda_1 > 1$ in Corollary 1 requires a positive mean correlation of returns, which is not a strong condition in practical terms.[4] It is worth mentioning that centrality does not necessarily rank assets in the same way as the mean correlation (mean of the row or columns in Ω) does. It is straightforward to show that for a correlation matrix with equal off-diagonal entries, each security is given the same amount of centrality and mean correlation (the leading eigenvector is $\frac{1}{\sqrt{n}}1$). However, this association breaks as the dispersion in the distribution of Ω_{ij} increases.[5]

Our theoretical results are consistent with Pozzi et al. (2013) in establishing that optimal portfolio strategies should overweigh low-central securities and underweight high-central ones. Therefore, optimal investors attempt to benefit from diversification by avoiding the allocation of wealth towards assets that are central in the correlation-based network. However, we depart from Pozzi et al. (2013) in two respects. First, our measure of centrality is derived from the investor's optimization problem, and thus is

[4] Rayleigh's inequality states a classical lower bound for the largest eigenvalues as follows $\lambda_1 \geq \frac{u'\Omega u}{u'u}$ for $u \in \mathbb{R}^n$. If $u = 1$ then $\lambda_1 \geq 1 + \frac{\sum_i \sum_j \Omega_{ij}}{n} = 1 + (n-1)\overline{\Omega}_{ij}$ where $\overline{\Omega}_{ij}$ is the mean correlation of off-diagonal elements of Ω.

[5] In a non-reported simulated exercise, a positive correlation between centrality and mean correlation is observed for a mild dispersion in the distribution of Ω_{ij}.

intimately associated with Markowitz's framework. Secondly, the individual performance of securities is overlooked in their study, leading to an incomplete analysis and potentially impairing the benefits of a network-based portfolio strategy.

3. Dataset description and stock market network estimation

In order to avoid data mining bias and to test the robustness of our results, four datasets are considered throughout the empirical sections accounting for different markets and time periods. Unless otherwise stated, split-and-divided-adjusted returns and traded volumes are obtained from CRSP while quarterly financial data comes from COMPUSTAT. The dataset *d-S&P* contains daily returns for 200 highly capitalized constituents of the S&P-500 index at the end of year 2012 showing non-negative total equity in the period Oct-2002 to Dec-2012. The dataset *m-NYSE* considers monthly returns for the 200 firms that remain listed in NYSE normalized for capitalization and cheap stocks in the period Jan-64 to Dec-06. [6] The dataset *d-FTSE* accounts for daily stock returns of the 200 most capitalized constituents of FTSE-250 index during the period Feb-06 to Oct-13. For this particular sample, we rely on Datastream as the data provider. A large dataset named *d-NYSE* is mainly used for simulation purposes and it considers daily returns for 947 firms listed in NYSE (adjusted for capitalization and cheap stocks explain as in *m-NYSE*) in the period Jan-2004 to Jul-2007. Finally, the risk-free rates required to compute excess returns for the US and UK markets are gathered from Kenneth French's website and from Gregory et al. (2013), respectively.

The large estimation error of the sample correlation matrix is well-documented (Jobson and Korkie, 1980). Therefore, we implement the shrinkage estimator the correlation matrix $\hat{\Omega}$ upon excess returns as in Ledoit and Wolf (2004) to estimate the adjacency matrix of the corresponding financial market network.

4. Fundamental drivers of stock centrality: descriptive analysis

Given the fundamental role assigned to the notion of centrality in this study, this section provides a set of descriptive results found in the *d-S&P* dataset. Table 1 reports the sample size, market capitalization and traded volume in 2012 (measured in millions of dollars) and the total and mean centrality by economic sectors (classified by firms' SIC codes). Although the manufacturing sector is the largest in terms of capitalization (48%) and traded volume (43%), financial firms are the most central nodes in the stock market network in accordance with reported evidence (Barigozzi and Brownlees, 2014; Peralta, 2015; Tse et al., 2010).

We also notice that the dispersion of the risk-adjusted returns' distribution is not constant across the network. The left panel of Fig. 1 plots the Sharpe ratio's boxplots conditioning on the low, middle and high terciles of the centrality distribution. Despite no significant difference in means, the Sharpe ratio's distribution shrinks as larger centralities are considered. Moreover, Table B.1 from Appendix B reports significantly greater mean volatility as we move from the bottom tercile (2.05%) to the middle (2.09%) and top tercile (2.24%) of the centrality distribution demonstrating larger risks among high-central securities. The relationship between the β-CAPM and the centrality of each security in the sample is plotted in right panel of Fig. 1.[7] This graph shows an upward-sloping relationship that, although not perfect, indicates that central assets tend to correspond to high systematic risk securities.

In order to identify the key financial and market drivers of stocks' centralities, we estimate an unbalanced quarterly-based panel regression largely inspired by Campbell et al. (2008) for the selection of relevant regressors. The study of Green and Hollifield (1992) rejects at a very high level of confidence the hypothesis that the first eigenvector of the covariance matrix is a constant vector and our econometric approach attempts to provide economic content to this result. A detailed description of the variables included in the regressions, the econometric techniques and results is provided in Appendix C. In a nutshell, our empirical analysis identifies central stocks with greater capitalization, lower price and older firms that maintain lower cash holdings. Therefore, we argue that the findings reported by Green and Hollifield (1992) can be explained in terms of the financial and market informational content that is embedded in the assets' centralities.

In an additional analysis reported in Appendix F to save space, we elaborate on the relationships between stocks' centrality and their stability. Specifically, we associate the concept of a stock's stability to the tendency to remain listed in the market without any change in its relative centrality status across time. We find that among the stocks that remain listed in the market, there is a strong tendency to show the same level of centrality through time. Additionally, we investigate the consequences of the period size used in the estimation of the correlation-based network on the ordering of securities provided by centrality. Our results show the large correlations between the rankings of centralities for different lengths of sample periods indicating the robustness of this ordering.

5. Stock centrality and optimal portfolio weights: in-sample evaluation

The interaction between the individual and the systemic dimensions of stocks is empirically investigated in this section. The dataset used is *d-S&P* and the results reported below come from both a cross-sectional and a time series in-sample analysis.

[6] The stocks are chosen to have capitalization more than 20th percentile of market capitalization and prices higher than \$5 (Penny stocks with prices lower than \$5 are discarded).

[7] The index S&P 500 is used as the market index for the estimation of the corresponding β from CAPM.

Table 1
Market capitalization, traded volume, and centrality by economic sectors.
This table reports the total and mean firms' centrality by economic sectors in the d-S&P500 dataset. Market capitalization and traded volume are in millions of dollars and correspond to the end of year 2012. The column Firms gives the number of companies in each economic sector. The columns denoted by % present the percentages with respect to the total of each of the preceding variables.

Economic sector	Firms	%	Market cap.	%	Traded vol.	%	Centrality	
							Total	Mean
Finance, insurance, and r. estate	37	19%	2,254,158	20%	70,788	19%	2.74	0.0741
Mining	17	9%	656,688	6%	26,178	7%	1.23	0.0722
Transp., comm., elect, gas, sanit. s.	27	14%	1,145,092	10%	36,988	10%	1.88	0.0696
Manufacturing	87	44%	5,338,584	48%	156,881	43%	6.01	0.0691
Retail trade	12	6%	745,486	7%	27,130	7%	0.82	0.0682
Services	13	7%	829,671	7%	37,314	10%	0.89	0.0682
Wholesale trade	5	3%	142,423	1%	6011	2%	0.33	0.0656
Construction	2	1%	44,905	0%	3073	1%	0.13	0.0646
Total	200		11,157,008		364,364			

5.1. Cross-sectional approach

The detailed pattern of co-movements across stocks is properly captured by $\hat{\Omega}$. This matrix conveys an excessive amount of information and leads to a fully connected stock market network that is difficult to analyze.[8] The Minimum Spanning Trees (MST), first introduced in the financial markets by Mantegna (1999), allows us to filter out this adjacency matrix with the aim of uncovering the market skeleton.[9] This technique has been widely applied to several country-specific stock markets including the US (Bonanno et al., 2004; Onnela et al., 2003), Korean (Jung et al., 2006), Greek (Garas and Argyrakis, 2007) and Chinese (Huang et al., 2009) markets among others.

Fig. 2 plots the MST financial market network for our data where nodes are scaled to the optimal portfolio weights for the *minv* strategy (see Eq. (3)) while the colors account for the respective security's centrality (darker colors imply greater centrality).[10] Note that investor's wealth is allocated toward lighter nodes (low-central securities) in accordance with Corollary 1. See the Internet Appendix for the full list of stocks with their respective centrality.

The relation between the individual and systemic dimensions of assets is illustrated in Fig. 3. The horizontal and vertical axes of this plot account for the securities' centralities and the corresponding standard deviations of returns, respectively. Each security is represented by a bubble whose size and color are given by the optimal portfolio weight in the *minv* rule. Note that most of the investor's wealth is assigned toward stocks located in the bottom-left corner of the graph, thus overweighting low-central-&-low-volatile securities.

Fig. 4 presents a scatter plot of Sharpe ratios and centralities for a portfolio applying *mv* strategy as the investment rule. The left and right panel sets the expected portfolio return, R^e, equal to 10% and 40% of the maximum possible portfolio return, respectively (see Eq. (6)). As expected, the optimal portfolio for low R^e mainly comprises assets with middle-ranged Sharpe ratios while the investment set moves toward securities with higher Sharpe ratios for larger R^e. Note, however, that the *mv* strategy avoids the allocation of wealth towards high-central stocks, say stocks with $v_i > 0.08$.

To further disentangle the roles of stock dimensions in the determination of optimal portfolio weights, we report in Table 2 the results from the OLS estimations of regressions (10) and (11). Eq. (10) considers the case of the *minv* strategy where stock *i*'s portfolio weight in a minimum-variance specification, $w_{i,minv}^*$, depends linearly on its centrality, v_i, and on its standard deviation of returns, σ_i. Similarly, Eq. (11) specifies a linear relationship between stock *i*'s portfolio weight in a mean-variance specification, $w_{i,mv}^*$, with the corresponding stock centrality, v_i, and Sharpe Ratio, SR_i.[11] Therefore, the coefficients β_1 and β_2 stands for the effects of the systemic and individual dimensions of stocks upon optimal portfolio weights.

$$w_{i,minv}^* = \beta_0 + \beta_1 v_i + \beta_2 \sigma_i + \varepsilon_i \tag{10}$$

$$w_{i,mv}^* = \beta_0 + \beta_1 v_i + \beta_2 SR_i + \varepsilon_i \tag{11}$$

The coefficients reported in Table 2 are strongly statistically significant and provide support to Proposition 1 and Corollary 1. The coefficient β_1 is negative for both regressions indicating that highly central stocks tend to be underweighted regardless of the specific investment objective. The coefficient β_2 is negative for the *minv* rule and positive for the *mv* case indicating the tendency to optimally allocate wealth towards low-standard deviation and high-Sharpe ratio securities, respectively.

[8] A fully connected network refers to a network structure in which each node is connected with the rest.
[9] MST connects the n stocks in a tree-like network by considering the highest $n-1$ paired correlations of returns as links to the extent that no loops are created.
[10] Short sales are allowed in the computation of optimal weights for both *minv* and *mv* strategies.
[11] In Eq. Eq.11, optimal weights are obtained assuming R^e equals 40% of the maximum possible return.

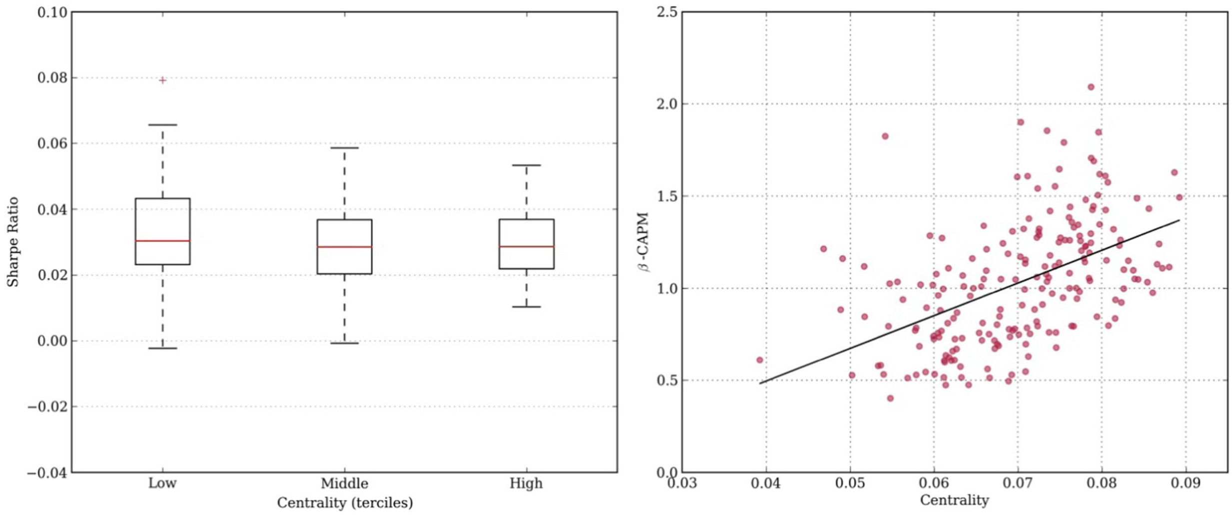

Fig. 1. Sharpe ratio distributions for the high, middle and low terciles of securities' centrality (left panel). Relationship between the securities' centrality and the β from CAPM (right panel). Both panels consider the d-S&P dataset.

5.2. Time series approach

Contrary to the static description provided by the cross-sectional analysis, this subsection takes a dynamic perspective on the portfolio selection by implementing a time series approach. The entire dataset is divided into 2522 60-day-long rolling windows considering 1-day displacement steps. The vectors of stock centrality, v_t and portfolio weights for the *minv* and *mv* rules, $w_{minv,t}^*$ and $w_{mv,t}^*$, are computed for each rolling window and indexed by the time subscript t.

We introduce the mean stock centrality $\overline{v}_t$ and the weighted stock centrality $\overline{\overline{v}}_{r,t}$ as follows:

$$\overline{v}_t = \frac{1}{n}\sum_{i=1}^{n} v_{it} \tag{12}$$

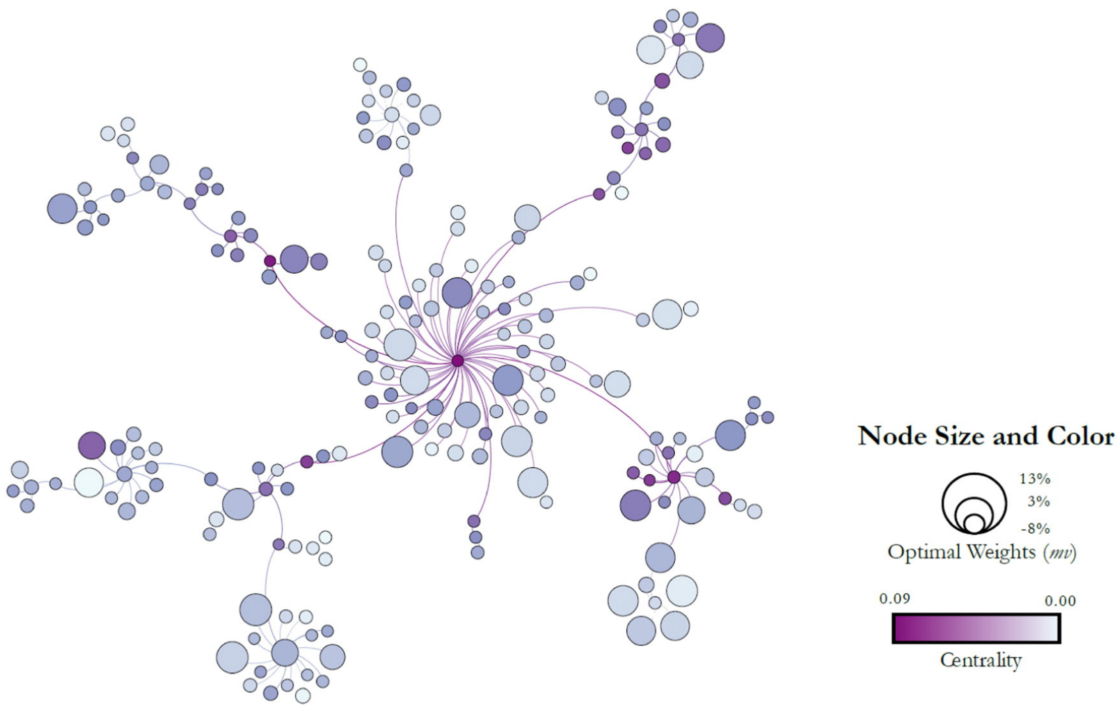

Fig. 2. MST stock market network for the d-S&P dataset. The size of nodes corresponds to the optimal weights for the minimum-variance strategy (see Eq. 3) and the intensity of the colour accounts for the corresponding security's centrality.

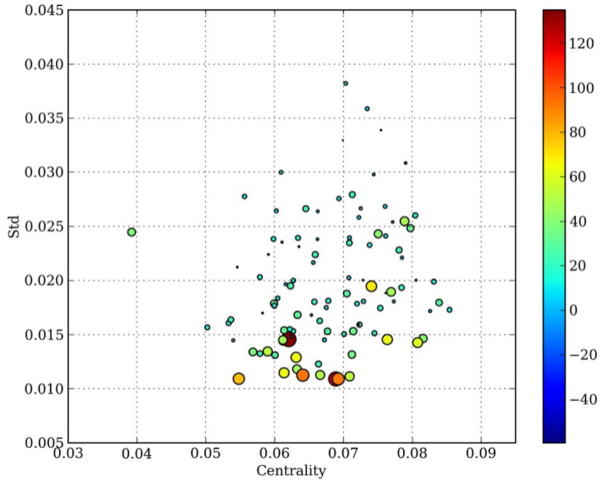

Fig. 3. Relationship between standard deviation of returns (Std) and stocks' centrality. Stocks correspond to bubbles whose sizes and colours (colour bar) reflect their optimal portfolio weights in a minimum-variance (*minv*) weights specification. We use d-S&P dataset for this analysis.

$$\overline{\overline{v}}_{r,t} = \sum_{i=1}^{n} w^{*}_{r,it}\, v_{it} \;\; for\; r = minv,\;\; mv \tag{13}$$

By construction, the expression $\overline{\overline{v}}_{r,t} = \overline{v}_t$for $r = minv,\;\; mv$ is satisfied when $w^{*}_{r,it} = 1/n$. In accordance with Corollary 1, it would be expected that $\overline{\overline{v}}_{r,t} < \overline{v}_t$ most of the time indicating a bias toward low-central stocks.

Fig. 5 plots the time series of $\overline{v}_t$ and $\overline{\overline{v}}_{r,t}$ presenting the pattern just described. Note, however, that $\overline{\overline{v}}_{r,t}$ shows values closer to, or even surpassing, $\overline{v}_t$ for some periods. The dynamic of the correlation between the individual and the systemic dimensions of stocks explains this time-dependent behavior of $\overline{\overline{v}}_{r,t}$. Let us denote by π_t the cross sectional correlation between v_{it} and σ_{it} and by ρ_t the cross-sectional correlation between v_{it} and SR_{it}, both at period t. When $\pi_t > 0$, the lowest standard deviation stocks tend to coincide with the weakly systemic ones, and as a consequence, overweighting these securities is certainly the optimal choice under the *minv* rule. For $\pi_t < 0$, a trade-off arises since low-central stocks also correspond to high-volatility securities. In this case, an optimal portfolios rule should balance these two confronting forces by adapting the investment set to include more central stocks. Considering the *mv* strategy, for $\rho_t < 0$, the assets with highest Sharpe ratios show the lowest centrality, and therefore, this leads to investing in non-systemic securities as the optimal portfolio choice. When $\rho_t > 0$, a trade-off between assets' dimensions takes

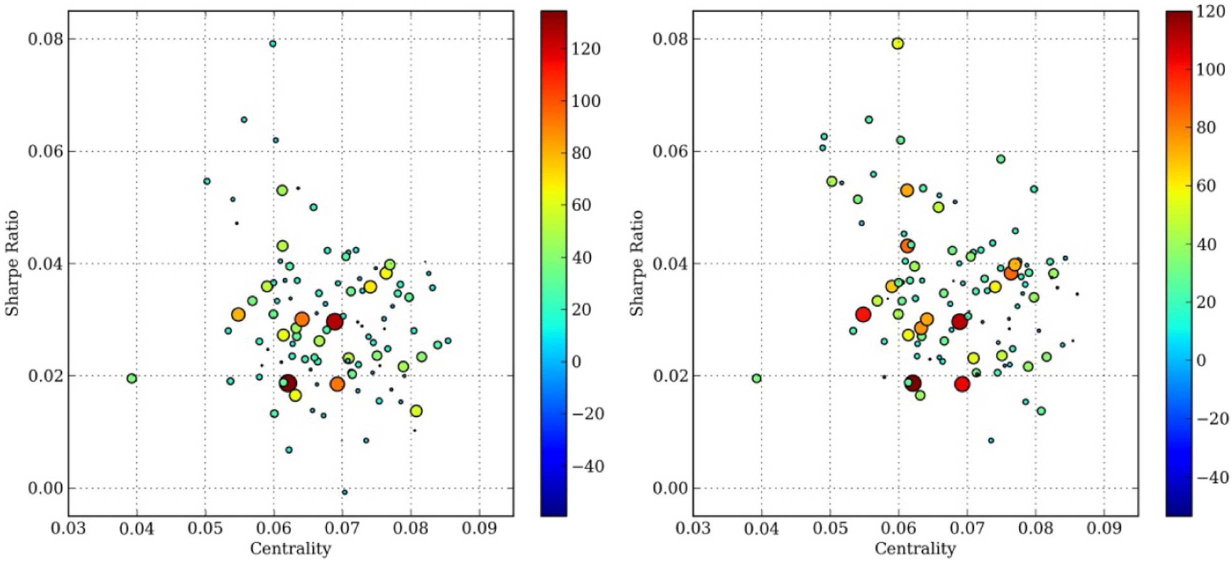

Fig. 4. Relationship between Sharpe ratio and stock centrality. Stocks correspond to bubbles whose sizes and colours (colourbar) reflect their optimal portfolio weights in a mean-variance (*mv*) weights specification. The required expected portfolio return R^e in Eq. 6 is 10% (left panel) and 40% (right panel).of the maximum return in the dataset. We used the d-S&P dataset for this analysis.

Table 2
Optimal portfolio weights as a function of the individual and systemic stock's dimensions considering a cross-sectional approach v_i is the centrality of stock i.. σ_i is the standard deviation of stock i. SR_i is the Sharpe ratio of stock i. N is the number of observations (stocks) in the cross-sectional regressions. The regression R^2 is adjusted for degrees of freedom. Each row of the table reports OLS estimations of Eqs. (10) and (11), respectively, where *t-statistics* are in parentheses.*, **

	v_i	σ_i	SR_i	N	R^2
$w^*_{i,minv}$	-0.740 (-4.05)***	-1.755 (-6.38)***		200	0.231
$w^*_{i,mv}$	-0.778 (-3.64)***		0.761 (4.77)***	200	0.179

* At 5% level.
** At 1% level.
*** At 0.1% level.

place and an optimal wealth allocation should increase portfolio weights towards central securities. Fig. 6 plots the time series of ρ_t and π_t and the corresponding 120-days moving averages evidencing the sign-switching nature of these two variables.

Further insights on the security selection process are gained by estimating the regressions (14) and (15) accounting for the time series versions of expressions (10) and (11). In (14), the optimally weighted centrality $\overline{\overline{v}}_{minv,t}$ for the *minv* strategy in period t is explained by the mean centrality, $\overline{v}_t$, the coefficient of variation of the centrality distribution, $(\frac{\sigma_{v,t}}{\overline{v}_t})$, and π_t. Similarly, Eq. (15) considers the *mv* strategy and therefore, the dependent variable is $\overline{\overline{v}}_{mv,t}$ while ρ_t replaces π_t as explanatory variable.

$$\overline{\overline{v}}_{minv,t} = \beta_0 + \beta_1\overline{v}_t + \beta_2\left(\frac{\sigma_{v,t}}{\overline{v}_t}\right) + \beta_3\pi_t + \varepsilon_t \tag{14}$$

$$\overline{\overline{v}}_{mv,t} = \beta_0 + \beta_1\overline{v}_t + \beta_2\left(\frac{\sigma_{v,t}}{\overline{v}_t}\right) + \beta_3\rho_t + \varepsilon_t \tag{15}$$

Table 3 reports OLS estimations of (14) and (15) noting that all of the coefficients are strongly statistically significant. The coefficient β_1 is negative for both regressions indicating that higher $\overline{v}_t$ results in an overweighting of low-systemic assets as a mean to avoid the undesirable consequences of high-central securities in the portfolio. The negative signs of β_2 are interpreted as the benefits derived from a wider centrality distribution that allows an increased presence of assets with lower centrality scores in the portfolio after controlling for $\overline{v}_t$. The coefficient β_3 is negative for the *minv* rule (see Eq. (14)). In this case therefore, larger values of π, indicating no trade-off between assets' dimensions, are consistent with optimal wealth allocations away from central securities. In contrast, the coefficient β_3 shows a positive sign for the *mv* rule (see Eq. (15)) explained by the trade-off between assets' dimensions arising for large ρ. As a consequence the increments of ρ rise $\overline{\overline{v}}_{mv}$ by moving the investment set toward central nodes.

6. Stock centrality and optimal portfolio weights: Out-of-sample evaluation

In DeMiguel et al. (2009), naïve strategy, commonly termed as $1/N$[12], is shown not to be consistently outperformed by Markowitz-based rules or their extensions designed to deal with the estimation error problem. This better out-of-sample performance of $1/N$ strategy relative to Markowitz's rule is also investigated and supported by Jobson and Korkie (1980); Michaud (2008) and Duchin and Levy (2009). Moreover, DeMiguel et al. (2009), p. 1936) report that among those models designed to tackle the estimation error problem, the constrained Markowitz rules might be considered as second-best alternatives to naïve diversification. Accordingly, naïve strategy and constrained Markowitz rules portray two reasonable benchmarks for out-of-sample portfolio evaluations.

Based on the insights obtained mainly from Sections 2 and 5, we here propose a network-based investment strategy, termed as *ρ-dependent strategy*, that targets groups of assets in accordance with their centrality rankings. We proceed to define this strategy in detail and evaluate its out-of-sample performance against the benchmarks.

6.1. The out-of-sample evaluation

The out-of-sample returns are computed as follows. For a T-period-long dataset, we implement several portfolio strategies (specified below) upon a set of M-period-long rolling windows indexed by the subscript t. We buy-and-hold the portfolios for H periods and the resulting out-of-sample return is recorded. The rolling window in period t is created by simultaneously adding the next H data points and discarding the H earliest ones from the previous rolling window in order to preserve its length. This process is repeated several times until the end of the dataset is reached thus accounting for $\lfloor (T-M)/H \rfloor$ rebalancing periods, vectors of portfolio's weights and out-of-sample returns for each portfolio strategy.

[12] Naïve strategy assigns a fraction $1/N$ of wealth to each asset out of the N available assets.

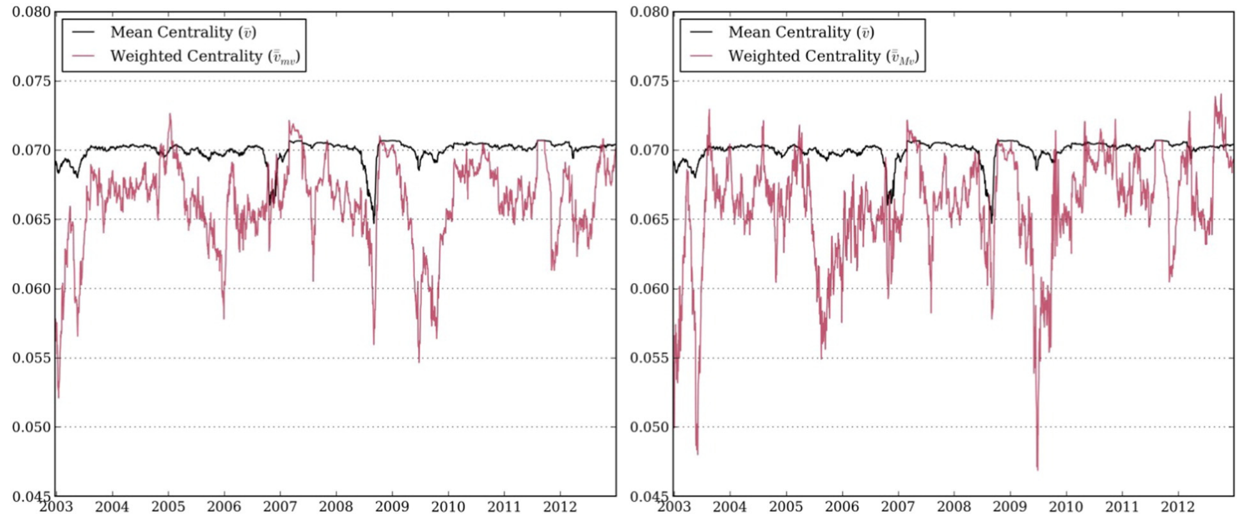

Fig. 5. Mean centrality $\overline{v}_t$ and weighted centrality $\overline{\overline{v}}_{r,t}$ through time. We consider the *minv* strategy (left panel) and the *mv* strategy (right panel) for the d-S&P dataset. We divide the dataset into 2522 of 60-days long rolling windows with 1-day displacement steps. The weight allocations of *minv* and *mv* strategies and the centrality rankings of stocks are computed for each rolling window. The mean centrality $\overline{v}_t$ is the average of centralities among stocks in each 60-days window and the weighted centrality, $\overline{\overline{v}}_{r,t}$, is computed by weighting each stock's centrality by their corresponding weights in *mv* and *minv* strategies.

We introduce our network based investment policy, the so-called *ρ-dependent strategy*, as follows. The process estimates the correlation matrix $\hat{\Omega}_t$, the vector of securities' centrality, v_t and the correlation between the systemic and individual dimensions of stocks, ρ_t, upon the rolling window corresponding to period t. Assuming a threshold parameter $\tilde{\rho}$, for a sufficiently large ρ, say $\rho > \tilde{\rho}$, we naively invest in the 20 stocks with the highest centrality. Conversely, for a sufficiently low ρ, say $\rho < \tilde{\rho}$, we naively invest in the 20 stocks with the lowest centrality. This strategy is designed to benefit from investing in low systemic stocks to the extent that the most central ones do not show significant individual performances, thus replicating Markowitz's logic to some extent.

In accordance with the previous literature (DeMiguel et al., 2009), the $1/N$ rule applied upon the entire investment opportunity set is a convenient benchmark with which to evaluate the performance of our network-based strategy. To ensure that the performance of ρ-dependent strategy does not happen by chance, we also includes the *reverse ρ–-dependent strategy* that naively invests in the 20 lowest central stocks when $\rho > \tilde{\rho}$ and in the 20 highest central ones when $\rho < \tilde{\rho}$. Moreover, two Markowitz-related rules, the mean-variance and minimum-variance strategies with short-selling constraints are incorporated in the analysis as well.

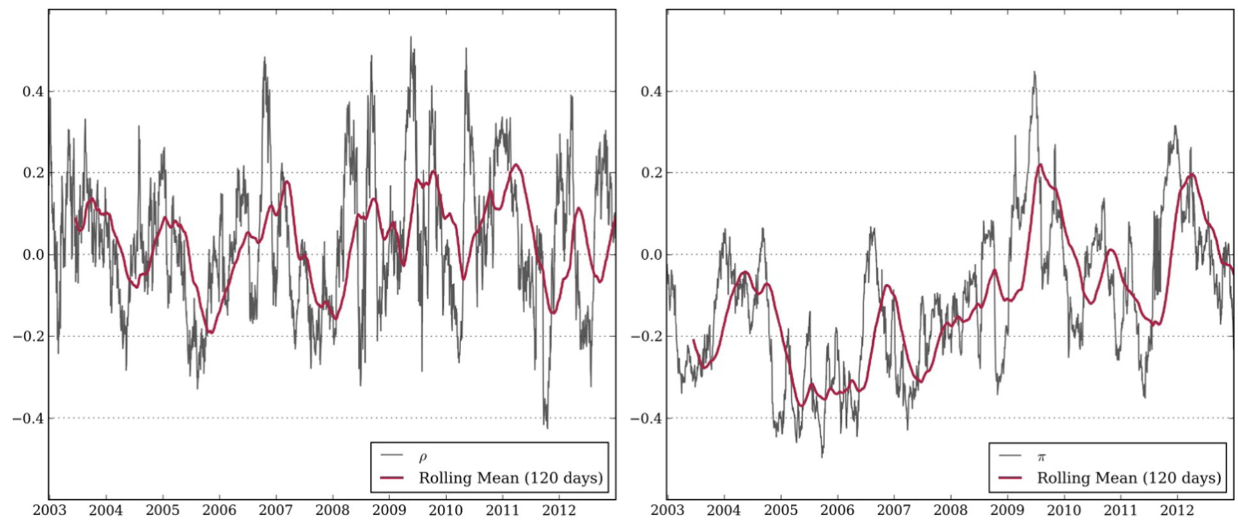

Fig. 6. Time series of the cross-sectional correlations between the individual and the systemic dimensions of stocks. This figure presents the correlation between stocks' centralities and Sharpe ratios, $\rho_t = corr(v_{it}, SR_{it})$ (left panel) and stocks' centralities and standard deviation $\pi_t = corr(v_{it}, \sigma_{it})$ (right panel) for the d-S&P dataset. We also include the corresponding 120-day moving averages of ρ_t and π_t in each panel.]

Table 3
Optimal portfolio weights as a function of the individual and systemic stock dimensions taking a time series approach We consider a 60-day rolling window estimation procedure for each variable. Each t denotes a 60-day rolling window. $\overline{v}_t$ is the mean centralities at each t. $(\frac{\sigma_{v,t}}{\overline{v}_t})$ is the coefficient of variation of the centrality distribution at t. π_t is the correlation between centralities, v_{it} and standard deviations, σ_{it} at t. ρ_t is the cross-sectional correlation between centralities, v_{it}, and Sharpe ratios, SR_{it}, at t. N is the number of observations (stocks) in the cross-sectional regressions. The regression R^2 is adjusted for degrees of freedom. Each row of the table reports OLS estimation of Eqs. (14) and (15). *t-statistics* are in parentheses.*, **

	$\overline{v}_t$	$\frac{\sigma_{v,t}}{\overline{v}_t}$	π_t	ρ_t	N	R^2
$\overline{\overline{v}}_{minv,t}$	-3.324 (-19.04)***	-0.0707 (-33.80)***	-0.00613 (-26.34)***		2522	0.573
$\overline{\overline{v}}_{mv,t}$	-2.661 (-10.83)***	-0.0720 (-24.50)***		0.00627 (17.79)***	2522	0.457

* At 5% level.
** At 1% level.
*** At 0.1% level.

Finally, we consider three more investment rules as "control strategies" that naively invest in the 20 stocks with the Highest Sharpe Ratio, the Highest Centrality and the Lowest Centrality[13].

The investment policies are compared using three out-of-sample performance measures: i) Sharpe ratio, ii) variance of return and iii) turnover. This latter measure averages the amount and size of the rebalancing operations as follows

$$turnover = \frac{1}{T-M-1}\sum_{t=M}^{T-1}\sum_{j=1}^{N}|w_{j,t+1}-w_{j,t^+}| \quad (16)$$

where $w_{j,t+1}$ is the weight of security j at the beginning of period $t+1$ and w_{j,t^+} is the weight of that security just before the rebalancing occurring between the end of period t and the beginning of period $t+1$.[14]

We statistically test the difference in out-of-sample portfolios' Sharpe ratios and variances between each of the investment rules against the $1/N$ strategy following Ledoit and Wolf (2008) and Ledoit and Wolf (2011). More specifically, we implement a studentized circular block bootstrap with block size equal to 5 and bootstrap samples equal to 5.000 to compute the respective p-values.[15]

6.2. Determining the threshold parameter $\tilde{\rho}$

Before the application of our network-based strategies, the value of $\tilde{\rho}$ needs to be specified. In order to do that, we rely on an extensive simulation procedure using artificially created subsamples where ρ spans a broad range of values. Specifically, we generate 120 datasets by randomly selecting 150 stocks from the d-NYSE dataset (see Section 3) with ρ ranging from -0.20 to 0.45. Then, we compute the out-of-sample Sharpe ratio where M and H are set to be equal to 500 days and 20 days, respectively.

Panels a, b and c from Fig. 7 plot the out-of-sample Sharpe ratios from three particular datasets with ρ equal to 0.45, -0.20 and 0.0, arising from a progressive percentage removal of stocks from the upper and lower tails of the centrality distribution (with 5% increments). Panel c from Fig. 7 (when $\rho=0$) presents a scenario in which an intensive removal of high central stocks leads to better out-of-sample performance in terms of Sharpe ratio. This result is expected since highly individual performing stocks are randomly disseminated in the stock market network. Panel (b) from Fig. 7 (when $\rho=-0.20$) shows us the convenience of selecting among low-central stocks as the target region to invest. This is due to the absence of trade-off between the systemic and individual dimension of stocks. The opposite case is observed in panel (a) of Fig. 7 (when $\rho=0.45$). In this case, the best performance is achieved by investing in high central stocks with an intense removal of securities from the left tail of the centrality distribution. This is explained given that the positive effect of the individual dimension of stocks over-compensates the negative effect of their systemic dimension.

In the next step, we investigate the break point of ρ that characterizes the region of the market network that is susceptible to being discarded from the investment opportunity set. Let us define the high central stock investment region as the set of securities arising from the deletion of 25% to 45% of assets from the left tail of the centrality distribution (starting from the lowest central security) and no more than 20% from its right tail (starting from the highest central security). In a symmetric fashion, let us define the low central stock investment region as the set of stocks comprising the deletion of 25% to 45% of stock from the right tail of the centrality distribution and no more than 20% of its left tail. Then, from all of the 120 artificially-constructed data sets, we identify the investment region that generates the highest out-of-sample Sharpe ratio. The identification rule simply averages

[13] The decision to construct portfolios made up of 20 assets is based on Desmoulins-Lebeault and Kharoubi-Rakotomalala (2012) which highlights the fact that most diversification benefits are gained by investing in such a number of stocks.

[14] Note that in the expression (16), rebalancing is assumed in each investment period as if $H=1$. However, with $H>1$, we only consider the rebalancing periods in our *turnover* calculation.

[15] For the particular case of the variance, a stationary bootstrap is employed as in Politis and Romano (1994).

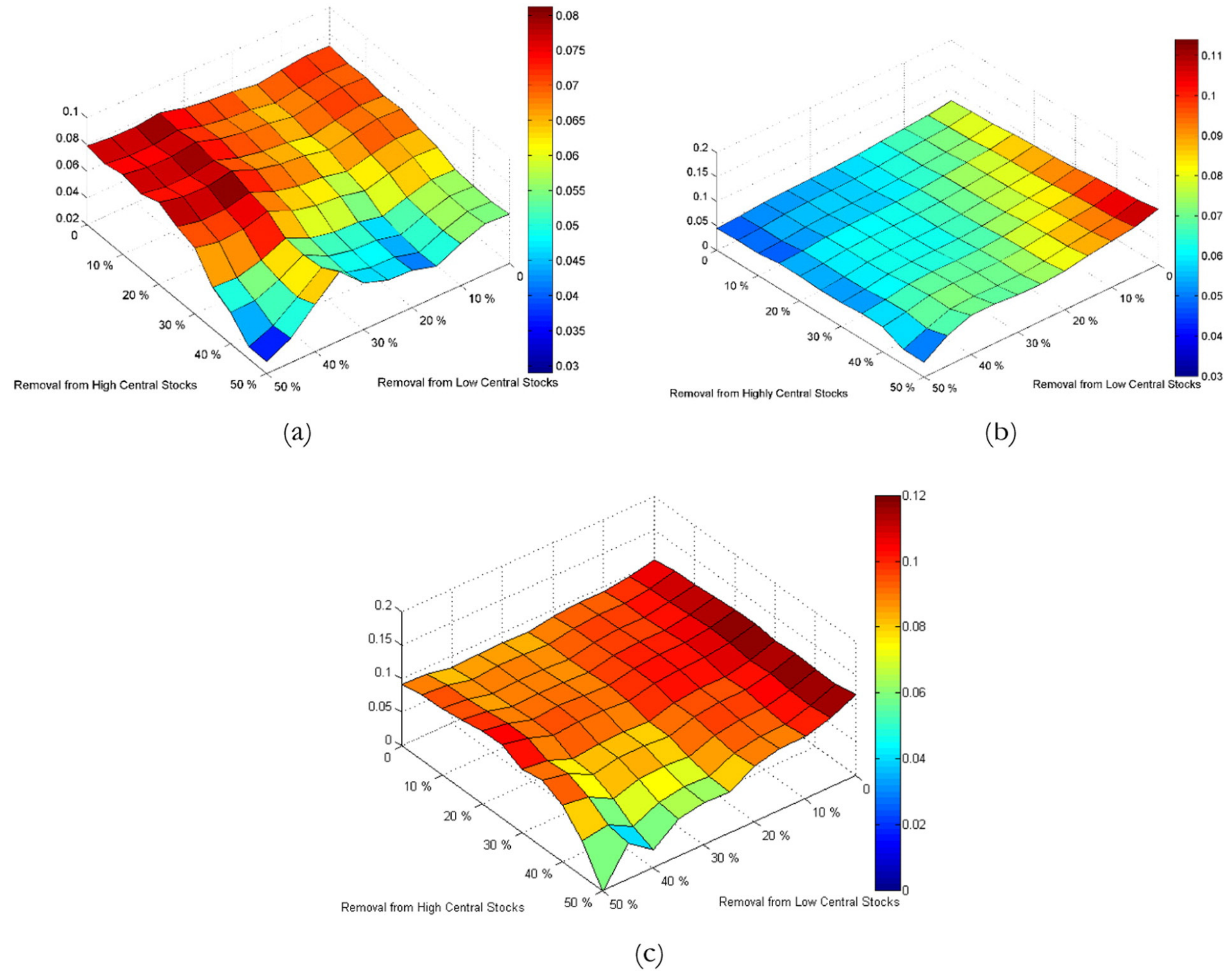

Fig. 7. Out-of-sample Sharpe ratios resulting from progressive percentage removal of stocks from the lower tail and upper tail of the centrality distribution. Three specific randomly generated samples from d-NYSE dataset are considered with ρ values: (a) $\rho = 0.4550$, (b) $\rho = -0.2010$, (c) $\rho = 0.000$. The out-of-sample Sharpe ratios are computed with $M = 500$ and $H = 20$.

out the out-of-sample Sharpe ratios generated in each of the two investment regions and then selects the one with the highest average performance. Fig. 8 plots the distribution of ρ conditional on the investment region that leads to the largest Sharpe ratio. In accordance with Fig. 8, low values of ρ are more consistent with high Sharpe ratio emerging from the low central investment region. In contrast, for large values of ρ, it is the high central stock investment region that generates the largest risk-adjusted returns.

With the aim of setting the value of $\tilde{\rho}$, note that such a threshold must be high enough to be worthwhile to move the optimal investment region from the low central securities towards more central ones.[16] Taking Fig. 8 into account, we consider 0.2 as a reasonable value for $\tilde{\rho}$ since it roughly coincides with the 75% percentile of ρ from the low central stock investment region and the 25% percentile of ρ from the high central stock investing region. We acknowledge that the determination of $\tilde{\rho}$ is the weakest point in our procedure since it implies ad-hoc rules. Other methodologies might be investigated in this regard and we leave them as future research lines.

6.3. Out-of-sample performance

The annualized out-of-sample performance measures originated from the two daily datasets d-S&P and d-FTSE and the monthly dataset m-NYSE are reported in Table 4. Since the average ρs from these datasets are lower than $\tilde{\rho} = 0.20$, the performance of the ρ-dependent strategy is indistinguishable from an unconditional lowest-central investment rule. Therefore, we also include an additional sample (d-NYSE150) obtained by a random selection of 150 stocks from d-NYSE presenting an average ρ equal to 0.23.

[16] The tendency is to invest in low-central stocks unless high-central ones show good individual performances.

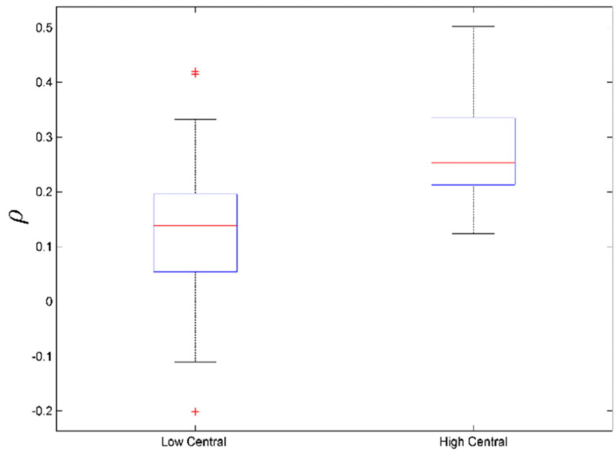

Fig. 8. Distribution of ρ conditioning on the investment region generating the highest out-of-sample Sharpe ratio. 120 artificially-created datasets from d-NYSE dataset is considered. *High Central* refers to the investment region comprised of the set of stocks arising from the deletion of 25% to 45% of assets from the left tail of the centrality distribution and no more than 20% from its right tail. *Low Central* refers to the investment region with the set of stocks comprising the deletion of 25% to 45% of stocks from the right tail of the centrality distribution and no more than 20% of its left tail. The out-of-sample Sharpe ratios are computed with $M = 500$ and $H = 20$.

The setup for the out-of-sample evaluation assumes $M = 1000$ and $H = 20$ for the daily datasets and $M = 192$ and $H = 12$ for the monthly dataset.[17]

Table 4 shows the noticeable outperformance of the ρ-dependent strategy for the daily datasets. The Sharpe ratio from the 1/N naïve strategy reaches 0.471 and 1.153 for d-S&P and d-FTSE, respectively. The same measures rise to 0.724 and 2.584 when the ρ-dependent strategy is in place, showing a statistically significant difference with respect to the benchmark. Similar good performances are observed in terms of portfolio variance, presenting a statistically significant reduction from 0.066 to 0.051 and from 0.025 to 0.014, for d-S&P and d-FTSE. Considering the m-NYSE dataset, the ρ–-dependent results are worse in terms of Sharpe ratio and variance in comparison to the benchmark but not to a statistically significant extent as indicated by the p-values. The benefit derived from the switching nature of ρ-dependent rules is explored by means of the d-NYSE150 dataset. In this case, a higher and statistically different Sharpe ratio (2.083) is obtained by following the proposed investment strategy in comparison to the naïve (1.241) and lowest-central (1.942) rules.

Table 4 also shows that ρ–-dependent strategy implies distinctive and enhanced investment dynamics when it is compared to the unconditional strategies. In general terms, the Highest Sharpe Ratio, Highest Central Stocks and Lowest Central Stocks strategies present poorer outcomes in a portfolio's Sharpe Ratio in all of the datasets. In order to discard the possibility that our results were driven by chance, Table 4 also reports the results for the reverse ρ–-dependent strategy. In this case, none of the out-of-sample portfolio's Sharpe ratios show better results when compared either with the benchmark or with the ρ-dependent strategy. Moreover, in the d-NYSE150 dataset, we observe that the ρ–-dependent rule significantly outperforms other strategies. Since the mean value of ρ in this dataset surpasses the threshold 0.2, the ρ-dependent strategy can exploit the benefit of moving the investment set from high-central to low-central nodes and vice versa.

Additionally, Table 4 reports the performance of the two Markowitz-related strategies: mean-variance and mean-variance with short selling constraints denoted by *minv-cc* and *mv-cc*, respectively. In comparison with these strategies, the ρ-dependent rule shows a better performance except for the case of *minv-cc* in the m-NYSE. For example, in the case of d-S&P dataset, while the *minv-cc* strategy manages to earn a lower level of portfolio variance compared to the benchmark, it fails to improve the level of the Sharpe ratio. On the other hand, the *mv-cc* strategy results in a non-significant increase in the Sharpe ratio level while maintaining the portfolio risk at the same level as the 1/N rule. In contrast, the ρ–-dependent strategy manages to simultaneously increase the level of the Sharpe ratio and decrease the portfolio variance level as commented previously. This comparison highlights the poor out-of-sample performance of Markowitz rules reported in the literature and also shows the convenience of implementing the network-based portfolio strategy.[18]

Unfortunately, the major shortcoming of our proposed rule is that it tends to raise the number of rebalancing operations, a phenomenon captured by the increased portfolio turnover compared to the 1/N benchmark. The severity of this issue is further investigated in Section 6.5 below. Two additional comments are worth mentioning. First, the relatively extreme and negative value of ρ (-0.1853 on average) for the d-FTSE dataset might explain the strikingly good results obtained in the UK market where no trade-off between the individual dimension and systemic dimension exists. Secondly, it should be mentioned that the prescription from Pozzi et al. (2013) in favor of unconditional allocation of wealth only towards the periphery of the stock market network shows clearly inferior results relative to the ρ-dependent rule when ρ assumes large values as in the d-

[17] Due to data limitation, $M = 500$ for the d-NYSE dataset.

[18] We also evaluate the performance of mean-variance and minimum-variance strategies without the short-selling constraints. However, our main conclusions remain unchanged. These results are available upon request.

Table 4
Out-of-sample performance of portfolio strategies.
We report the out-of-sample Sharpe ratio, variance and turnover for portfolio strategies. The benchmark strategy is denoted as "*All stocks*", that is naïve strategies applied to all of the stocks in the dataset. Following the *ρ-dependent* strategy, when ρ is higher than 0.2, we diversify among highest central stocks and otherwise, diversify among the lowest central stocks. The *Reverse ρ-dependent* approach takes an opposite investment decision to the ρ-dependent strategy. The *Highest Sharpe Ratio* strategy refers to diversifying naively among stocks with the highest level of Sharpe ratios. In *Lowest Central* and *Highest Central* strategies, we diversifying among lowest and highest central stocks, respectively. The *minv-cc* and *mv-cc* strategies refer to minimum-variance and mean-variance strategies with short-selling constraints. We considered 20 stocks for our portfolios. The p-values are computed following the procedure in Ledoit and Wolf (2008) and Ledoit and Wolf (2011) based on a studentized circular block bootstrap with block size equal to 5 and number of bootstrap samples equal to 5000.

Panel A	d-S&P (Avg ρ: -0.0556)			d-FTSE (Avg ρ: -0.1853)		
	Sharpe Ratio	Variance	Turnover	Sharpe Ratio	Variance	Turnover
All Stocks	0.471	0.066	0.141	1.153	0.025	0.138
ρ-dependent	0.724	0.051	0.149	2.523	0.014	0.164
	(0.0125)	(0.0010)		(0.0033)	(0.0009)	
Reverse ρ-dependent	0.315	0.110	0.126	0.549	0.026	0.061
	(0.1523)	(0.0009)		(0.0100)	(0.014)	
Highest Sharpe Ratio	0.438	0.038	0.141	1.201	0.018	0.138
	(0.8239)	(0.0009)		(0.8007)	(0.0009)	
Highest Central	0.315	0.110	0.126	0.549	0.026	0.061
	(0.1523)	(0.0009)		(0.0100)	(0.014)	
Lowest Central	0.724	0.051	0.149	2.523	0.014	0.164
	(0.0125)	(0.0010)		(0.0033)	(0.0009)	
minv_cc	0.448	0.040	0.115	1.589	0.015	0.127
	(0.9568)	(0.0020)		(0.5482)	(0.0010)	
mv-cc	0.573	0.067	0.148	1.089	0.025	0.137
	(0.8339)	(0.8951)		(0.9535)	(0.6563)	
Panel B	**m-NYSE (Avg ρ: 0.1157)**			**d-NYSE150 (Avg ρ: 0.2323)**		
	Sharpe Ratio	Variance	Turnover	Sharpe Ratio	Variance	Turnover
All Stocks	0.754	0.019	0.057	1.241	0.021	0.127
ρ-dependent	0.591	0.025	0.065	2.083	0.019	0.140
	(0.3997)	(0.2553)		(0.0664)	(0.3367)	
Reverse ρ-dependent	0.552	0.034	0.049	0.968	0.028	0.113
	(0.0233)	(0.0009)		(0.5216)	(0.0020)	
Highest Sharpe Ratio	0.531	0.017	0.039	1.0571	0.018	0.103
	(0.1694)	(0.2248)		(0.5781)	(0.0110)	
Highest Central	0.555	0.032	0.049	1.201	0.033	0.106
	(0.0299)	(0.0009)		(0.8671)	(0.0009)	
Lowest Central	0.583	0.028	0.065	1.942	0.014	0.147
	(0.0831)	(0.0009)		(0.1927)	(0.0009)	
minv_cc	0.777	0.015	0.046	1.310	0.013	0.109
	(0.9302)	(0.1558)		(0.9734)	(0.0110)	
mv-cc	0.456	0.054	0.053	1.353	0.039	0.122
	(0.3422)	(0.0010)		(0.9502)	(0.0010)	

NYSE150 dataset. Nevertheless, in the cases where ρ is relatively low, as in the d-S&P and d-FTSE datasets, our results are in line with theirs.

6.4. Carhart alpha for the ρ-dependent strategy

The large risk-adjusted returns of the ρ-dependent strategy might result from large exposures to systematic risk factors. We investigate this hypothesis by estimating Carhart's alpha from the four risk factor models (Carhart, 1997; Fama and French, 1996, 1993) as in Eq. (17).

$$R_t^\rho - r_t^f = \alpha + \beta_{MKT}\left(R_t^M - r_t^f\right) + \beta_{HML}HML_t + \beta_{SMB}SMB_t + \beta_{MOM}MOM_t + \varepsilon_t \quad (17)$$

where $R_t^\rho - r_t^f$ is the excess out-of-sample return from the ρ-dependent strategy, $R_t^M - r_t^f$ is the market risk premium, HML_t is the difference between the returns of high and low book-to-market portfolios, SMB_t is the difference between the returns of small-cap and large-cap portfolios and MOM_t is the momentum factor. The parameter α measures the abnormal risk-adjusted return capturing the excess return above what would be expected based solely on the portfolio's risk profile. The time series of the four risk factors are gathered from Ken French's website for the US market and from (Gregory et al., 2013) for the UK case. We correct the standard errors in the estimation of Eq. (17) for heteroskedasticity and serial correlation in ε_t following Newey and West (1987).

Table 5 reports the estimated Carhart's alpha for three portfolio sizes (10, 20 and 50) and two holdings periods (1 and 20 days for the daily datasets and 1 and 12 months for the monthly dataset) across the US and UK samples. The reader is referred to

Appendix D for a detailed results of the estimation of Eq. (17) without winsorizing. However, to account for outliers, Table 5 reports results after 5% data winsorizing, noting that qualitative similar results are obtained either without winsorizing or by winsorizing at 10% (see Tables E.1 and E.2 in Appendix E). The estimations show positive and statistically significant alphas for each of the portfolio configurations. The weakest results stem from the monthly dataset where the reduced estimation window might undermine the statistical significance.

To sum up, the ρ-dependent rule is able to provide enhanced risk-adjusted returns that are not explained by exposure to traditional risk factors. Interestingly, the evidence also indicates a negative relationship between portfolio sizes and alphas for each of the considered holding periods and across datasets. For instance, considering the d-S&P dataset and a 20-day holding period, the Carhart's alphas of portfolios made of 10, 20 and 50 stocks are 17.19, 11.97 and 10.08, respectively. To confirm this regularity, however, further analyses are required.

Finally, Fig. 9 plots the cumulative out-of-sample returns of the ρ-dependent strategy for different portfolio configurations. For the d-S&P dataset plotted in panel (a), we observe that investing in 10 stocks results in the largest payoffs while the $1/N$ rule produces the poorest performance. The same observation applies for the d-FTSE dataset captured in panel (b) from the same figure. The case of the m-NYSE dataset presented in panel (c) is different; the time series of cumulative returns of the naïve strategy outperforms the ρ-dependent strategy for portfolios made of 10 and 20 stocks while for a portfolio size equal to 50 they both tend to behave similarly.

6.5. Transaction cost

Since our ρ-dependent strategy is a dynamic strategy that requires rebalancing operations, it is important to investigate the impact of transaction costs. Following the approach provided in Han et al. (2013), we compute the breakeven transaction cost (BETC) that sets the average returns of the ρ-dependent strategy equal to zero. Denoting by R_t the out-of-sample return of the ρ-dependent strategy in period t, the breakeven transaction cost (BETC) is computed as follows:

$$BETC = \frac{\sum_{t=M+1}^{T} R_t}{T-M} \times H \tag{18}$$

As stated in Balduzzi and Lynch (1999), reasonable lower and upper bounds for the transition cost are 1 bp and 50 bps, respectively. Thus, obtaining a BETC higher than 50 bps means that a disproportionally abnormal transaction cost is required to wipe out the returns of the ρ-dependent strategy.

The results for BETC are reported in Table 6 considering the d-S&P, d-FTSE and m-NYSE datasets and different portfolio settings (size and holding period). In general terms, we observe that for short holding periods, BETC assumes low values for any investment configuration. For instance, in a setting of a 1 day holding period and 10 stocks from the FTSE dataset, a transaction cost of only 13.79 bps is needed to wipe out the benefit of ρ-dependent strategy. However, for longer holding periods, our calculations indicate large values of BETC which supports the outperformance of the ρ-dependent strategy after controlling for transaction cost. For example, considering again the FTSE dataset, a portfolio size of 10 and a holding period of 20 days, a transaction cost of more than 264.56 bps is needed to wipe out the returns of the proposed strategy.

It should be noted that in this analysis, it is assumed that the investor faces a flat-transaction fee every time he/she wants to rebalance the portfolio regardless the size of the rebalancing operation. Appendix G provides a detailed analysis in a context in which the investor pays in proportion to the changes in portfolio's weights. The results reported in this appendix provide further support to the application of our proposed network-based investment rule.

Table 5
Annualized risk-adjusted returns for ρ-dependent strategy with 5% winsorisation.
We report annualised risk-adjusted returns for different settings of ρ-dependent strategy on the four Carhart (1997) factors, MKT, HML, SMB, MOM with 5% winsorisation. The estimation window is considered to be 1000 days (192 months) for daily (monthly) datasets. The t-statistics are reported in parentheses.

Portfolio setting		Alpha (5% winsorizing)		
Stocks	Holding period	d-S&P	d-FTSE	m-NYSE
10	1d/1m	14.75	35.97	5.1
		(3.37)**	(6.37)**	(2.26)*
	20d/12m	17.19	36.01	4.49
		(3.97)**	(6.21)**	(1.98)*
20	1d/1m	10.57	33.65	4.42
		(3.02)**	(6.65)**	(2.31)*
	20d/12m	11.97	32.84	4.37
		(3.45)**	(6.52)**	(2.36)*
50	1d/1m	10.52	28.07	5.45
		(3.63)**	(5.17)**	(3.23)**
	20d/12m	10.08	28.08	5.82
		(3.50)**	(5.18)**	(3.40)**

** Indicate significance at the 1% level.
* Indicate significance at the 5% level.

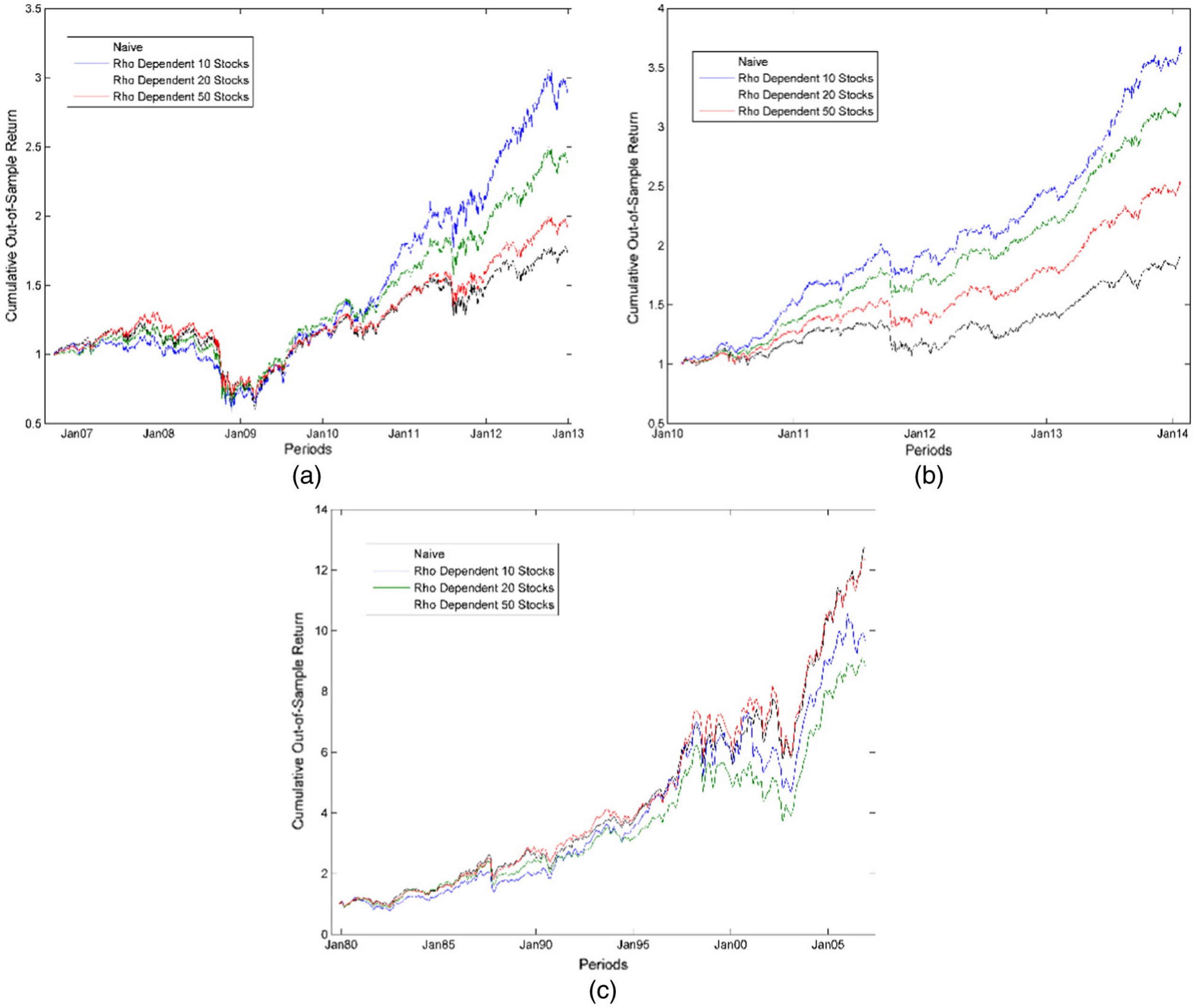

Fig. 9. Cumulative return for the ρ-dependent strategy for (a) d-S&P, (b) d-FTSE, (c) m-NYSE datasets in comparison to the naïve strategy for 20 days and 12 month holding periods in daily and monthly datasets, respectively.

7. Conclusion and future research lines

In this study, a stock market is conceived as a network where securities correspond to nodes while the paired returns' correlations account for the links. The paper establishes a bridge between Markowitz's framework and network theory by stating the tendency to overweight low-central stocks in order to build efficient portfolios. Therefore, optimal portfolio weights of highly influential securities in a correlation-based network are biased downward after controlling for their individual performance measured by either Sharpe ratios or the volatility of returns (depending on the specific portfolio's goal).

From a more descriptive point of view, we find that financial firms are the most central nodes in the market network and that both financial and market variables are major determinants of a stock's centrality. More precisely, we provide some evidence indicating that highly central securities correspond to large-capitalized, cheap and old firms presenting lower cash holdings.

We also investigate the extent to which network-based investment strategies might improve portfolio performance by means of in-sample and out-of-sample analysis. We propose the so-called ρ-dependent strategy and test its performances against the extremely simple yet effective $1/N$ naïve rule and two Markowitz-related policies. Our out-of-sample results show that the ρ-dependent strategy tends to present significant higher portfolio Sharpe ratios and lower portfolio variance relative to these well-known benchmarks. Additionally, this enhanced performance is not explained by large exposures to traditional risk factors as indicated by the reported positive and statistically significant Carhart's alphas. More importantly, our results are robust to several portfolio configurations, time periods and markets even after accounting for transaction costs.

There are several future research lines that could provide novel insights from the interaction between network theory and portfolio selection. However, it seems particularly appealing to extend the approach by considering stock markets as *directed* *and weighted* networks. This framework may contribute to improve our understanding on the shock-transmission mechanisms across stocks and disentangling the differential role played by specific securities as an absorber or booster of initial impulses.

Acknowledgements

We thank Francisco J. Nogales, Pedro Serrano, David Moreno, Ricardo Crisóstomo and the participants in the 22nd Finance Forum of the Spanish Finance Association (AEFIN) and in the internal seminars at both CNMV and Universidad Carlos III de Madrid for helpful comments. We are also grateful to the editor, Daniel Ferreira, and the anonymous reviewer for their insightful suggestions to improve the manuscript. The opinions in this paper are the sole responsibility of the authors and they do not necessarily coincide with those of the CNMV. Gustavo Peralta acknowledges financial support from the CNMV.

Appendix A. Proof of Proposition 1

Let us consider two symmetric square $n\times n$ matrices Ω^0 and $\Omega^1=\Omega^0+a^*I$ where $a\in\mathbb{R}$ and I is the identity matrix with their corresponding sets of eigenvectors and eigenvalues denoted by $\{v_1^s,\ldots,v_n^s\}$ and $\{\lambda_1^s,\ldots,\lambda_n^s\}$ for $s=1,0$. By definition of eigenvectors $\lambda_k^s v_k^s=\Omega^s v_k^s$, it follows that $\lambda_k^1 v_k^1=\Omega^1 v_k^1=(\Omega^0+a^*I)\ v_k^1$. After some simple algebraic manipulations $(\lambda_k^1-a)v_k^1=\Omega^0 v_k^1$ that allows us to conclude that $v_k^1=v_k^0$ and $\lambda_k^1=\lambda_k^0+a$. Therefore, as a preliminary result, we show that the eigenvectors of Ω^0 and $\Omega^1$are exactly equal and the corresponding associated eigenvalues are related as follows: $\lambda_k^1=\lambda_k^0+a$ for $k=1\ldots n$.

The proof of Proposition 1 is stated only for the case of the mean-variance strategy given that the minimum-variance rule follows exactly the same steps. We assume that the correlation matrix Ω is a $n\times n$ diagonalizable symmetric matrix with a set of eigenvectors given by $\{v_1,\ldots v_n\}$ and a set of eigenvalues given by $\{\lambda_1,\ldots\lambda_n\}$, both sets arranged in descendent order. Then, $\Omega=P\Lambda P^T$, where P is an $n\times n$ orthogonal matrix whose columns are $v_1,\ldots v_n$. Let us denote by $\Lambda=diag(\lambda_i)$ a diagonal matrix whose *ith*-main diagonal element is λ_i. Thus the inverse of Ω could be written as

$$\Omega^{-1}=P\Lambda^{-1}P^T=P*diag(1/\lambda_i)*P^T \tag{A.1}$$

$$\Omega^{-1}=\sum_k\left(\frac{1}{\lambda_k}v_k{v_k}^T\right)=\frac{1}{\lambda_1}v_1{v_1}^T+\frac{1}{\lambda_2}v_2{v_2}^T+\ldots+\frac{1}{\lambda_n}v_n{v_n}^T \tag{A.2}$$

From Eq. (7) in Section 2.2, we have $\hat{w}^*=\varphi\Omega^{-1}\hat{\mu}^e$. By adding and subtracting $\varphi\hat{\mu}^e$ from this expression we get

$$\hat{w}^*=\varphi\hat{\mu}^e+\varphi\left[\Omega^{-1}-Isssssssss\right]\hat{\mu}^e \tag{A.3}$$

Using the preliminary results above-mentioned in this appendix, we know that the matrix $\Omega^{-1}-I$ has the same eigenvectors with eigenvalues equal to $\frac{1}{\lambda_k}-1$ for $k=1\ldots n$. Therefore, A.3 is stated as follows:

$$\hat{w}^*=\varphi\hat{\mu}^e+\varphi\left[\sum_k\left(\frac{1}{\lambda_k}-1\right)v_k{v_k}^T\right]\hat{\mu}^e \tag{A.4}$$

Given that eigenvector centralities refer to the elements of the eigenvector corresponding to the largest eigenvalue, we define $\Gamma=\varphi[\sum_{k=2}^n(\frac{1}{\lambda_k}-1)v_k{v_k}^T]\hat{\mu}^e$. Then, A.4 is stated as

$$\hat{w}^*=\varphi\hat{\mu}^e+\varphi\left(\frac{1}{\lambda_1}-1\right)v_1{v_1}^T\hat{\mu}^e+\Gamma \tag{A.5}$$

$$\hat{w}^*=\varphi\hat{\mu}^e+\varphi\left(\frac{1}{\lambda_1}-1\right)\left({v_1}^T\hat{\mu}^e\right)v_1+\Gamma \tag{A.6}$$

Table 6
BETC for ρ-dependent strategy.
We report the BETC for the ρ–-dependent strategy in various strategy settings across the three datasets of d-S&P, d-FTSE and m-NYSE. The results are reported in basis points (bps).

Portfolio setting		BETC		
Stocks	Holding period	d-S&P	d-FTSE	m-NYSE
10	1d/1m	6.98	13.79	85.69
	20d/12m	156.74	264.56	937.18
20	1d/1m	5.42	12.12	79.69
	20d/12m	130.1	236.59	982.12
50	1d/1m	5.04	9.67	85.5
	20d/12m	102.06	191.76	1040.94

$$\hat{w}^* = \varphi\hat{\mu}^e + \varphi\left(\frac{1}{\lambda_1} - 1\right)\hat{\mu}_M^e v_1 + \Gamma \tag{A.7}$$

Appendix B. Descriptive statistics of stock performance in terms of centrality

Table B.1
Sharpe ratios, excess returns and standard deviations of returns for the stocks included in the d-S&P dataset conditioning on their centralities. The stocks are categorised in terms of the low, middle and high terciles of the centrality distribution.

	Mean	Std	Percentiles					Skew	Kurtosis
			Min	5%	50%	95%	Max		
Sharpe ratio									
Low	3.31%	1.57%	−0.22%	1.23%	3.03%	6.16%	7.92%	0.51	0.13
Middle	2.82%	1.13%	−0.07%	0.88%	2.85%	4.35%	5.86%	−0.02	−0.06
High	2.88%	0.92%	1.03%	1.36%	2.86%	4.04%	5.33%	0.00	−0.66
Excess return									
Low	0.07%	0.05%	−0.01%	0.02%	0.06%	0.18%	0.22%	1.20	0.84
Middle	0.06%	0.03%	0.00%	0.02%	0.06%	0.12%	0.13%	0.58	−0.08
High	0.06%	0.02%	0.02%	0.03%	0.06%	0.09%	0.16%	0.85	2.59
Std									
Low	2.05%	0.70%	1.09%	1.20%	1.95%	3.21%	4.84%	1.30	2.53
Middle	2.09%	0.63%	1.09%	1.15%	2.03%	3.24%	3.82%	0.58	−0.23
High	2.24%	0.50%	1.43%	1.49%	2.23%	3.17%	3.76%	0.60	0.23

Appendix C. A panel regression for the stock centrality

In order to identify the key financial and market drivers of stocks' centralities, we estimate an unbalanced quarterly-based panel regression largely inspired by Campbell et al. (2008) for the selection of relevant regressors. Our empirical exercise attempts to provide an economic content to the result reported in Green and Hollifield (1992). As before, the dataset used is d-S&P dataset comprising 7931 firms-quarter data points.

The dependent variable is the centrality of a security (firm) i in quarter t.[19] The financial explanatory variables are ROA_{it}, Lev_{it} and Liq_{it} accounting for the ratio of *net income*, *total liability* and *cash and short term assets* to *total assets*, respectively. As market explanatory variables, we include $\ln(MV_{it})$ and $\ln(P_{it})$, denoting the logarithms of the market capitalization on a common-shares basis and the stocks' market price at the end of quarter t, respectively. In addition, $\ln(TV_{it})$, Ret_{it} and Std_{it}, referring to the logarithm of total trades, the excess return and the daily standard deviation of returns during quarter t, are additional market explanatory variables. The variable M/B_{it} denotes the Market-to-Book ratio on a common-shares basis at the end of quarter t and is incorporated in the regressions as well. Finally, the logarithm of firms' age indicated as $\ln(Age_{it})$ comprises the last independent variable of the model. It is computed by counting the number of quarters elapsed since the appearance of the first market price in CRSP until period t as in Fama and French (2004). To control for the effects of outliers, 1% winsorizing is implemented on regressors except for the case of $\ln(Age_{it})$. Table C.1 reports summary statistics.

Table C.2 presents OLS estimations for two specifications of the panel regression described above. Model I includes dummy variables by quarter and economic sector (first 2 digits of SIC codes) and considers only robust-heteroskedastic standard errors (White, 1980)[20]. To tackle the bias induced by autocorrelation and heteroskedasticity in the error term, we also report estimations of model II that considers economic sector dummy variables while clustering the error term by quarters as suggested by Petersen (2009).

Starting with the analysis of financial explanatory variables, Table C.2 indicates that the coefficients of *ROA* and *M*/*B* are negative in each specification, however, they are not significant at conventional levels. Similarly, the variable *Lev* stays positive but non-significant across models. Finally, the coefficient for variable *Liq* is negative and strongly significant in models I and II. Therefore, there is some evidence indicating that firms' centrality conveys information on the financial risk profile of the sampled firms.

Among market explanatory variables, ln(*MV*) shows positive and strongly significant coefficients across specifications evidencing a size effect. The variables ln(*TV*) and ln(*P*) present a negative impact on centrality for each specification and remain significant for both models. Therefore, low-traded and cheaper securities tend to be highly central in the financial market network. The variable *Ret* is positive for each model but statistically significant only for the first model. In the case of *Std*, since it is marginally significant only for model I and changes its sign across specifications, we disregard its effects. Finally, ln(*Age*) shows positive coefficients that are strongly significant for models I and II.[21]

[19] Only the data corresponding to quarter t is considered for computing the centralities. Additionally, we rescaled centrality by multiplying it by 100 for exposition purposes.

[20] We also estimate model I with two further specifications (i) with industry and time fixed effect while applying two-ways clustering, and (ii) with time and firm fixed effects. We find no significant changes in the results.

[21] We are grateful to the anonymous referee for suggesting the inclusion of ln(*Age*) in the analysis.

Table C.1
Descriptive statistics for the quarterly panel regression variables included in the d-S&P dataset.
The description of the variables is as follows: *ROA* is net income/total assets at the end of period *t*, *Lev* is total liability/total assets at the end of period *t*, *Liq* is cash and short term assets over the total assets at the end of period *t*, *ln*(*MV*) is the logarithm of market capitalization on a common-share basis at the end of period *t*, *ln*(*TV*) is the logarithm of total trades during period *t*, *Ret* is the excess return during period *t*, *ln*(*P*) is the logarithm of stocks' prices at the end of period *t*, *Std* is the return variance during period *t*, *M/B* is the market-to-book ratio on a common-share basis at the end of period *t* and *ln*(*Age*) is the logarithm of the firms' age computed as in Fama and French (2004) and corresponding to the end of period *t*. The descriptive statistics are reported after 1% winsorising except for *ln*(*Age*).

Variables	Mean	Std	Percentiles					Skew	Kurtosis
			Min	5%	50%	95%	Max		
Independent									
Financial									
ROA	1.70%	1.70%	−5.00%	−0.10%	1.50%	4.70%	7.50%	0.10	2.70
Lev	59.00%	19.50%	12.20%	23.20%	59.60%	90.80%	94.50%	−0.20	−0.40
Liq	12.10%	13.40%	0.20%	0.70%	6.90%	41.20%	67.00%	1.90	3.70
Market									
ln(*MV*)	16.98	1.02	14.20	15.34	16.88	18.91	19.42	0.10	0.20
ln(*TV*)	12.47	1.04	9.95	10.83	12.41	14.28	15.21	0.20	0.10
Ret	2.60%	13.60%	−35.10%	−20.50%	2.80%	25.00%	45.10%	0.10	1.00
ln(*P*)	3.69	0.58	2.23	2.72	3.70	4.58	5.54	0.10	0.60
Std	1.80%	1.10%	0.70%	0.80%	1.50%	3.90%	6.80%	2.40	6.90
M/B	3.32	2.50	0.59	0.95	2.57	8.50	14.71	2.11	5.42
ln(*Age*)	4.84	0.77	1.79	3.43	4.96	5.81	5.86	−0.66	−0.12
Dependent									
Centrality	7.01	0.93	4.01	5.27	7.11	8.41	8.95	−0.69	0.86

Table C.2
OLS estimations of three specifications of the quarterly panel regression model.
Each specification depends on the particular standard error correction method. *t-statistics* are in parentheses and the statistical significance is as follows: * at 5% level, ** at 1% and *** at 0.1% level. Model I combines economic sector and quarterly dummies with robust-heteroscedastic standard errors (White, 1980). Model II includes economic sector dummies while clustering standard errors by quarters. The regressions are implemented upon the d-S&P dataset.

	Models	
	I	II
ROA	−1.128	−1.062
	(−1.44)	(−1.03)
Lev	0.111	0.111
	(1.31)	(1.29)
Liq	−0.481	−0.582
	(−4.36)***	(−4.13)***
M/B	−0.00523	−0.00865
	(−0.90)	(−1.15)
ln(*MV*)	0.239	0.216
	(8.96)***	(8.32)***
ln(*TV*)	−0.230	−0.186
	(−7.88)***	(−6.10)***
Ret	0.401	0.196
	(4.23)***	(1.16)
ln(*P*)	−0.239	−0.168
	(−6.40)***	(−4.02)***
Std	−5.029	2.399
	(−2.50)*	(0.97)
ln(*Age*)	0.104	0.124
	(5.53)***	(6.78)***
N	7931	7931
R^2	0.174	0.165
Dummies	*Econ. Sectors and Quarter*	*Econ. Sectors*
Std. errors' correction	*Robust Heteroscedastic*	*Clustering by Quarters*

Appendix D. Regressions of returns from ρ-dependent strategy on the four risk factors, MKT, HML, SMB, MOM without winsorizing

Table D.1
OLS estimation of Annualized Carhart's Alpha from the four factor model in Eq. (17).We consider the d-S&P, and d-FTSE, m-NYSE datasets and without winsorising. *t-statistics* are in parentheses. Standard errors are corrected following Newey and West (1987). Portfolios of size 10, 20 and 50 are considered with a holding period of one (1d) and twenty (20d) for the daily datasets and one (1m) and twelve (12m) months for the monthly dataset.

Dataset	Obs.	Portfolio setting		α	β_{MKT}	β_{HML}	β_{SMB}	β_{MOM}
		Stocks	Holding period					
SP500	1580		1d/1m	12.87	0.87	−0.15	0.06	0.18
		10		(2.74)***	(53.93)***	(−3.32)***	(1.49)	(0.76)
			20d/12m	15.06	0.9	−0.17	0.01	0.04
				(3.20)***	(38.60)***	(−3.34)***	(0.16)	(1.72)*
			1d/1m	3.5	0.89	−0.19	−0.07	0.03
		20		(2.60)***	(54.95)***	(−5.28)***	(−2.47)**	(2.02)**
			20d/12m	11.79	0.9	−0.21	−0.1	0.03
				(3.33)***	(33.82)***	(−4.39)***	(−2.39)**	−1.61
		50	1d/1m	8.14	0.92	−0.14	−0.14	0.05
				(3.44)***	(61.96)***	(−5.06)***	(−5.93)***	(3.95)***
			20d/12m	8.27	0.93	−0.17	−0.17	0.04
				(3.47)***	(47.79)***	(−5.60)***	(−5.60)***	(3.57)
FTSE	1000		1d/1m	35.45	−0.01	0.19	−0.02	0.01
		10		(5.37)***	(−0.39)	(0.30)	(−0.39)	(0.14)
			20d/12m	34.03	−0.01	0.03	−0.03	0.02
				(5.07)***	(−0.07)	(0.51)	(−0.54)	(0.55)
			1d/1m	31.27	−0.2	0.04	0.01	0.03
		20		(5.13)***	(−0.61)	(0.73)	(0.18)	(0.85)
			20d/12m	30.49	−0.02	0.03	0.01	0.03
				(5.02)***	(−0.59)	(0.57)	(0.11)	(0.86)
			1d/1m	24.77	−0.02	0.02	0.03	0.06
		50		(3.77)***	(−0.54)	(0.32)	(0.6)	(1.37)
			20d/12m	24.59	−0.02	0.02	0.03	0.05
				(3.73)***	(−0.54)	(0.36)	(0.59)	(1.35)
NYSE	324		1d/1m	0.88	0.91	0.32	0.36	−0.07
		10		(0.38)	(17.16)***	(4.64)***	(4.31)***	(−0.38)
			20d/12m	1.4	0.89	0.41	0.32	−0.1
				(0.62)	(17.55)***	(4.81)***	(4.37)***	(−1.64)
			1d/1m	1.69	0.95	0.41	0.29	−0.14
		20		(0.93)	(23.08)***	(5.79)***	(4.91)	(−2.64)***
			20d/12m	1.11	0.93	0.38	0.27	−0.14
				(0.63)	(21.60)***	(5.63)***	(4.78)***	(−2.99)***
			1d/1m	2.01	0.91	0.38	0.23	−0.13
		50		(1.40)	(26.68)***	(5.59)***	(3.93)***	(−2.77)***
			20d/12m	2.25	0.89	0.38	0.24	−0.12
				(1.59)	(26.04)***	(5.85)***	(3.97)***	(−2.64)***

* At 5% level.
** At 1% level.
*** At 0.1% level.

Appendix E. Annualized risk-adjusted returns for ρ-dependent strategy with 0% and 10% winsorization

Table E.1
Annualized risk-adjusted returns for different settings of ρ-dependent strategy without winsorisation
We report annualized risk-adjusted returns for different settings of the ρ-dependent strategy on the four Carhart (1997) factors, MKT, HML, SMB, MOM. The estimation window is considered to be 1000 days (192 months) for daily (monthly) datasets. The t-statistics are reported in the parentheses.

Portfolio setting		Alpha		
Stocks	Holding period	S&P500	FTSE250	NYSE
10	1d/1m	12.87	35.45	0.88
		(2.74)***	(5.37)**	(0.38)
	20d/12m	15.06	34.03	1.4
		(3.20)**	(5.07)**	(0.62)
20	1d/1m	3.5	31.27	1.69
		(2.60)**	(5.13)**	(0.93)
	20d/12m	11.79	30.49	1.11
		(3.33)**	(5.02)**	(0.63)

Table E.1 *(continued)*

Portfolio setting		Alpha		
Stocks	Holding period	S&P500	FTSE250	NYSE
50	1d/1m	8.14 (3.44)**	24.77 (3.77)**	2.01 (1.4)
	20d/12m	8.27 (3.47)**	24.59 (3.73)**	2.25 (1.59)

** Indicate significance at 1% level.
* Indicate significance at 5% level.

Table E.2
Annualized risk-adjusted returns for ρ-dependent Strategy with 10% winsorisation.
We report annualized risk-adjusted returns for various settings of ρ-dependent strategy on the four Carhart (1997) factors, MKT, HML, SMB, MOM with 10% winsorisation. The estimation window is considered to be 1000 days (192 months) for daily (monthly) datasets. The t-statistics are reported in the parentheses.

Portfolio setting		Alpha		
Stocks	Holding period	10% winsorizing		
		S&P500	FTSE250	NYSE
10	1d/1m	14.81 (3.82)**	35.14 (7.06)**	5.16 (2.48)*
	20d/12m	17.85 (4.57)**	34.75 (6.88)**	5.3 (2.55)*
20	1d/1m	13.83 (4.24)**	34.47 (7.77)**	9.98 (4.29)**
	20d/12m	15.72 (4.86)**	33.7 (7.60)**	5.91 (3.39)**
50	1d/1m	12.68 (4.51)**	29.51 (6.19)**	5.78 (3.70)**
	20d/12m	13.14 (4.70)**	29.74 (6.70)**	6.45 (4.03)**

** Indicate significance at 1% level.
* Indicate significance at 5% level.

Appendix F. The relationship between stock centrality and stability

We associate the concept of stock stability with the tendency of a particular asset to remain listed in the market through time without any change in its relative centrality status. This appendix analyzes the stability of stocks listed in NYSE from two different perspectives. First, we present results regarding the switching nature of assets in accordance to their different centrality in the stock market network. Second, we investigate whether the period size chosen to compute the correlation matrix influence the ranking of centralities.

We employ an m-NYSE dataset that accounts for all of the NYSE stocks with monthly pricing records in the period starting from April-1968 until April-2012. Thus, we include a full list of companies that have ever existed at some point in this period. We analyze the change in the nature of stocks in terms of centrality by relying on a moving window approach. We specify a 30-year moving window and divide it into two sub-periods, each of 15 years. We use the first and the second period to give an initial and final categorization of stocks in accordance with their centrality. Three exclusive and collectively exhaustive groups of stocks are created: high, medium and low central stocks in accordance to the top, middle and bottom terciles of the centrality distribution. Next, we construct a switching matrix accounting for the distribution of stocks belonging to a particular range of centrality in the initial period, in terms of their centrality in the final period. Since a one-year displacement step is under consideration, 15 individual switching matrices were computed. Fig. F1 reports the results for each of those iterations[22].

[22] The sum of vertical height for each line does not sum to one since the proportion of the delisted firms is not included in the graph.

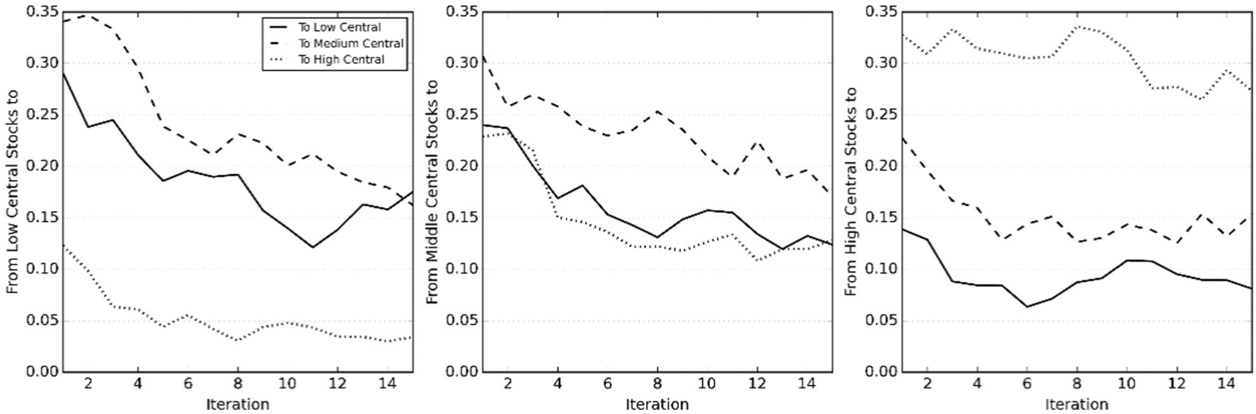

Fig. F.1 Change in centrality of securities from an initial tercile of centrality to a final tercile of centrality across 15 iterations in m-NYSE dataset. We specify a 30-year moving window and divide it into two sub-periods, each of 15 years. We use the first and the second period to have an initial and final categorization of stocks securities in accordance with their centrality. Three exclusive and collectively exhaustive groups of stocks are created: high, medium and low central stocks in accordance to the top, middle and bottom terciles of the centrality distribution.

In the left-panel of Fig. F1, it is obvious that most of low central stocks tend to change their nature to medium central or stay low central. Additionally, we can see that just a small proportion of low central stocks change to become highly central. The middle-panel of Fig. F1 depicts the changing nature of medium central stocks. As it can be inferred, medium central stocks mostly stay in the middle range of centrality and lower percentages of them tend to change towards the low or high centrality bucket. Finally, the right-panel of Fig. F1 presents the result of the experiment when stocks were initially classified as highly central assets. We can observe that high central stocks tend to stay central across the iterations and only lower percentage of them tend to become low or medium central. As a summary, Fig. F2 presents the average of the percentage of switching in centrality over the 15 interactions. In a nutshell, we find that among the stocks that remain listed, there is a tendency to occupy the same position into the stock market network through time providing a sort of stability to the structure.

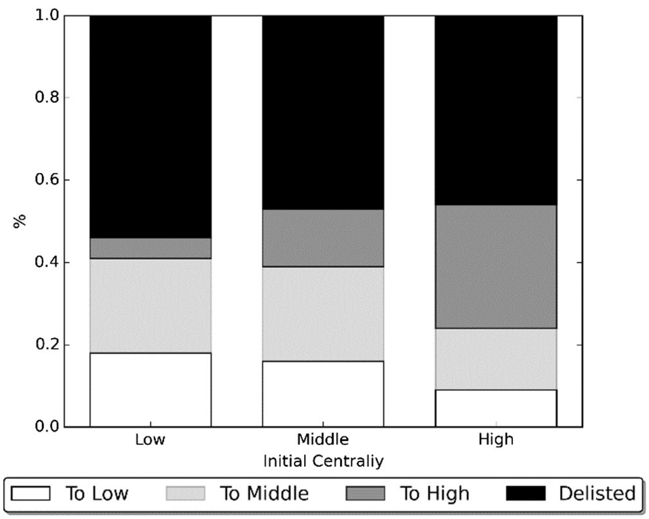

Fig. F.2 Average percentage of centrality switching for low, medium and high central securities. A 30-year moving window is specified in m-NYSE dataset and it is divided into two sub-periods, each of 15 years. We use the first and the second period to give an initial and final categorization of securities in accordance to their centrality. Three exclusive and collectively exhaustive groups of assets are created: high, medium and low central stocks, in accordance with the top, middle and bottom terciles of the centrality distribution.

In the second analysis, we investigate the relationship between the ranking of centrality and the size of the period chosen to compute the correlation matrix. We consider the d-S&P dataset for this analysis because several lengths of period can be investigated with a daily frequency.

We split the d-S&P data into several yearly subsamples made of 250 daily returns and compute the centrality rankings for each of these subsamples. In the next step, we further divide the yearly subsamples into shorter datasets accounting for one month (25 daily returns), two month (50 returns) and so on up to the yearly subsamples. Afterward, the mean correlations between the ranking of centralities obtained from these shorter datasets with the initial yearly subsamples are computed and reported in Fig. F.3. Clearly, the correlation between these centrality rankings is high (on average 0.8) indicating that the ordering it provides is considerably robust to the period size used to estimate the correlation matrix.

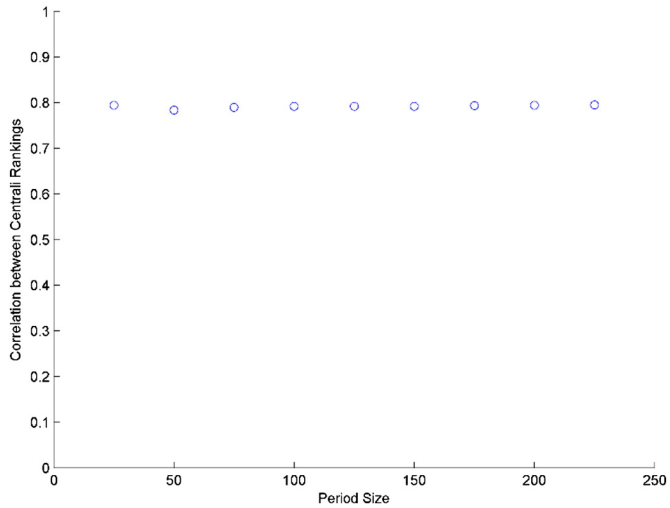

Fig. F.3. Average correlation of centrality rankings in the d-S&P500 dataset between yearly subsamples (250 returns) and a set of lower-size subsamples starting with length equal to 25 days and subsequent increments of 25 days up to 250 days (one year).

Appendix G. Turnover-driven transaction cost

The analysis in Section 6.5 assumes that investor faces a fixed transaction cost without accounting for the magnitude of changes in portfolio weights. However, investors may incur proportional costs that severely undermine the performance of their strategy. In order to take this point into account, we compute another version of the breakeven transaction cost (BETC) by solving the next expression:

$$\sum_{t=M}^{T-1}\left[(1+R_t)\left(1-BETC^{*}\times\sum_{j=1}^{N}\left|w^{S}_{j,t+1}-w_{j,t^{+}}\right|\right)-1\right]=0 \tag{G.1}$$

Therefore, the value of $BETC^{*}$represents the corresponding proportional of transaction cost that is required in order to eliminate portfolio return. The results are presented in Table G.1 where low values of $BETC^{*}$ are displayed. Nevertheless, considering the low proportional transaction cost that investors face in the current state of the financial markets, we conclude that the ρ-dependent strategy remains profitable even from this perspective.

Table G.1
$BETC^{*}$ for ρ-dependent strategy.
We report the $BETC^{*}$ for ρ–-dependent strategy in various strategy settings across the three datasets of d-S&P, d-FTSE and m-NYSE. The results are reported in basis points.

Portfolio setting		BETC		
Stocks	Holding period	d-S&P	d-FTSE	m-NYSE
10	20d/12m	13	20	12
20	20d/12m	12	19	14
50	20d/12m	11	18	13

Since the corresponding $BETC^{*}$s for one day/month holding period are extremely low, Table G.1 does not report their values. In these cases, even a small proportional transaction cost might eliminate any benefit which in turn raises the concern derived from the application of the ρ-dependent strategy in a highly frequent rebalancing framework.

Appendix H. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jempfin.2016.06.003.

References

Balduzzi, P., Lynch, A.W., 1999. Transaction costs and predictability: some utility cost calculations. J. Financ. Econ. 52, 47–78.

Barigozzi, M., Brownlees, C., 2014. Nets: Network Estimation for Time Series. SSRN - Work. Pap., pp. 1–43

Billio, M., Getmansky, M., Lo, A.W., Pelizzon, L., 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. J. Financ. Econ. 104, 535–559. http://dx.doi.org/10.1016/j.jfineco.2011.12.010.

Bloomfield, T., Leftwich, R., Long, J., 1977. Portfolio strategies and performance. J. Financ. Econ. 5, 201–218.

Bonacich, P., 1972. Factoring and weighting approaches to status scores and clique identification. J. Math. Sociol. 2, 113–120.

Bonacich, P., 1987. Power and centrality: a family of measures. Am. J. Sociol. 92, 1170–1182.

Bonanno, G., Caldarelli, G., Lillo, F., Miccichè, S., Vandewalle, N., Mantegna, R.N., 2004. Networks of equities in financial markets. Eur. Phys. J. B Condens. Matter 38, 363–371.

Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In Search of Distress Risk. J. Financ. 63, 2899–2939.

Carhart, M., 1997. On persistence in mutual fund performance. J. Financ. 52, 57–82.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy? Rev. Financ. Stud. 22, 1915–1953.

Desmoulins-Lebeault, F., Kharoubi-Rakotomalala, C., 2012. Non-Gaussian diversification: when size matters. J. Bank. Financ. 36, 1987–1996.

Diebold, F.X., Yilmaz, K., 2014. On the network topology of variance decompositions: measuring the connectedness of financial firms. J. Econ. 182, 119–134.

Duchin, R., Levy, H., 2009. Markowitz versus the Talmudic portfolio diversification strategies. J. Portf. Manag. 35, 71–74.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56.

Fama, E., French, K., 1996. Multifactor explanations of asset pricing anomalies. J. Financ. 51, 55–84.

Fama, E., French, K., 2004. New lists: fundamentals and survival rates. J. Financ. Econ. 73, 229–269. http://dx.doi.org/10.1016/j.jfineco.2003.04.001.

Freeman, L., 1978. Centrality in social networks conceptual clarification. Soc. Networks 1, 215–239.

Garas, A., Argyrakis, P., 2007. Correlation study of the Athens stock exchange. Physica A 380, 399–410.

Green, R., Hollifield, B., 1992. When will mean-variance efficient portfolios be well diversified? J. Financ. 47, 1785–1809.

Gregory, A., Tharyan, R., Christidis, A., 2013. Constructing and testing alternative versions of the Fama–French and Carhart models in the UK. J. Bus. Financ. Acc. 40, 172–214.

Han, Y., Yang, K., Zhou, G., 2013. A new anomaly: the cross-sectional profitability of technical analysis. J. Financ. Quant. Anal. 48, 1433–1461.

Hautsch, N., Schaumburg, J., Schienle, M., 2015. Financial network systemic risk contributions. Eur. Finan. Rev. 9, 685–738.

Huang, W.Q., Zhuang, X.T., Yao, S., 2009. A network analysis of the Chinese stock market. Physica A 388, 2956–2964.

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

Jobson, D.J., Korkie, B.M., 1980. Estimation for Markowitz efficient portfolios. J. Am. Stat. Assoc. 75, 544–554.

Jorion, P., 1991. Bayesian anc CAPM estimators of the means: implications for portfolio selection. J. Bank. Financ. 15, 717–727.

Jung, W., Chae, S., Yang, J.S., Moon, H., 2006. Characteristics of the Korean stock market correlations. Physica A 361, 263–271.

Laloux, L., Cizeau, P., Bouchaud, J.P., Potters, M., 1999. Noise dressing of financial correlation matrices. Phys. Rev. Lett. 83, 1467–1470. http://dx.doi.org/10.1103/PhysRevLett.83.1467.

Ledoit, O., Wolf, M., 2004. Honey, I shrunk the sample covariance matrix. J. Portf. Manag. 30, 110–119.

Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. J. Empir. Financ. 15, 850–859.

Ledoit, O., Wolf, M., 2011. Robust performance hypothesis testing with the variance. Wilmott 55, 86–89.

Mantegna, R.N., 1999. Hierarchical structure in financial markets. Eur. Phys. J. B - Condens. Matter 11, 193–197.

Markowitz, H., 1952. Portfolio selection. J. Financ. 7, 77–91.

Merton, R., 1980. On estimating the expected return on the Market. J. Financ. Econ. 8, 323–361.

Michaud, R., 2008. Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation. Second. ed. Oxford University Press, New York.

Michaud, R., Michaud, R.O., 2008. Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation. Second. ed. Oxford University Press, New York.

Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Newman, M.E.J., 2004. Analysis Of Weighted Networks. Phys. Rev. E.

Onnela, J.P., Chakraborti, A., Kaski, K., Kertész, J., Kanto, A., 2003. Asset trees and asset graphs in financial markets. Phys. Scr. T106, 48–54.

Peralta, G., 2015. Network-based measures as leading indicators of market instability: the case of the Spanish stock market. J. Netw. Theory Financ. 1, 91–122.

Petersen, M., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Rev. Financ. Stud. 22, 435–480.

Politis, N., Romano, J., 1994. The STATIONARY bootstrap. J. Am. Stat. Assoc. 89, 1303–1313.

Pozzi, F., Di Matteo, T., Aste, T., 2013. Spread of risk across financial markets: better to invest in the peripheries. Sci. Rep. 3, 1665.

Trzcinka, C., 1986. On the number of factors in the arbitrage pricing model. J. Financ. 41 (374–368).

Tse, C.K., Liu, J., Lau, F.C.M., 2010. A network perspective of the stock market. J. Empir. Financ. 17, 659–667.

Van Mieghem, P., 2011. Graph Spectra for Complex Networks. Cambridge University Press.

Vandewalle, N., Brisbois, F., Tordoir, X., 2001. Non-random topology of stock markets. Quant. Financ. 1, 372–374.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica.

Zareei, A., 2015. Network centrality, failure prediction and systemic risk. J. Netw. Theory Financ. 1.