

Complex Networks in Finance

Anna Maria D’Arcangelis and Giulia Rotundo

Abstract The present paper can be considered as divided in two parts: in the first one, we provide a review of the methods of complex networks that have been mainly used in the applications to the analysis of financial data. We focus on the following topics: the usage of the correlation matrix, systemic risk, integrated ownership and control, board of directors, interbank networks, and mutual funds holdings structure. The second part shows this last subject and provides new analyses.

The main findings outline that there are substantial differences in geographical allocation among the different European fund managers. Five larger European countries dominate the market of mutual funds. The belonging of UK and Swiss opt-outs of the eurozone could be a probable explanation for our results on community detection, that give a snapshot of a sort of “geographical organization” of the core of mutual funds portfolios.

Keywords Complex networks • Correlation matrix • Financial markets • Integrated ownership • Mutual funds • Systemic risk

1 Introduction

The nouns *graphs* and *networks* refer to the same abstract structure, although they are used in different scientific areas for different purposes. Indeed, networks became popular after the exploitation of social networks in the '30s and '50s (Borgatti et al. 2009), while the foundation of the formal building of a graph dates back to Euler, who first presented his results on the Königsberg bridges problem in 1735. The

A.M. D’Arcangelis (✉)

Department of Economics and Business, University of Tuscia, DEIM, via del Paradiso 47,
01100 Viterbo, Italy

e-mail: adarcangelis@unitus.it

G. Rotundo

Department of Methods and Models for Economics, Environment and Finance, La Sapienza
University of Rome, via del Castro Laurenziano 9, 00161 Rome, Italy

e-mail: giulia.rotundo@uniroma1.it

© Springer International Publishing Switzerland 2016
P. Commendatore et al. (eds.), *Complex Networks and Dynamics*,
Lecture Notes in Economics and Mathematical Systems 683,
DOI 10.1007/978-3-319-40803-3_9

209

river crossing the old Prussian city of Königsberg had two islands connected by seven bridges, and the problem was to find a path to make a complete tour of the city crossing each bridge only once.

The formalization proposed by Euler differed from classic geometrical problems because it did not involve distances of meters: the land was represented through circles (nodes, units, elements), and the bridges through lines (edges, arches, links). This abstraction highlights the main characteristics of the problems. The proof that was given of the impossibility of the existence of such a path is based on the count of the number of edges connecting each node (node degree). Still now, the detection of paths with specific features constitutes a relevant task in graph theory, and it has a wide range of practical applications, from the optimal design of databases and print of electronic circuits to the travelling salesman problem.

The approach outlined above differs from the methods of combinatorics, which date back to a few centuries before, and leads to probability theory. In the middle of the twentieth century, the insertion of the probability theory into graphs boosted a new field of studies, *random graphs*.

Paul Erdos and Alfréd Rényi’s famous cooperation generated a series of papers, the most well known of which introduced the Erdos-Rényi model of random graphs (Newman et al. 2006). Targets in random graph theory are the detection of the probability of the presence of a specific property (defined through a variable) in graphs drawn from a particular distribution. This constitutes a meeting point with problems rising from Physics, like the percolation theory that characterizes the connectedness of random graphs. Physicists were already familiar with regular graphs/networks (lattices), mostly in the framework of ferromagnetism and statistical mechanics. The main input for passing from “simple” networks to *complex networks* raised from the studies in social sciences (Albert and Barabasi 2002). Empirical data evidenced that, besides the randomness, the network was not showing a trivial structure, but revealed features that do not occur in simple networks. Hence, the term “complex”. Nowadays, complex networks constitute an active and promising area of scientific research, widely inspired by the empirical analysis of real-world networks. It is part of *network science*, coded by the United States National Research Council as “the study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena”.

Excellent reviews are available on the theory and applications of complex networks in several different areas (Albert and Barabasi 2002; Barrat et al. 2004; Boccaletti et al. 2006; Borgatti et al. 2009; Bougheas and Kirman 2014; Pastor-Satorras et al. 2003; Varela Cabo et al. 2015; Newman et al. 2006; Nature, focus issue 2013).

In financial markets research, literature records a fast growth in scientific production, mostly grounded in Physics, and an increase in the level of cross-disciplinary perspectives. First, concepts for the analysis of social networks found a proper representation, and, later, the literature recorded a burst of studies tackling research issues that can be conveniently managed through complex networks. Most of the studies have an empirical approach and provide a good base for the development of new mathematical models, econometric analysis, and open new perspectives

for understanding large-scale phenomena such as the contagion channels in the financial system, the interconnections among financial institutions and markets and the analysis of systemic risk and financial stability. In other words, the relevance and role of single elements in the network can be evidenced, and critical areas identified.

In this chapter, the discussion of the papers analysing financial data with methods of complex networks will be conducted into four sections: analysis of the correlation matrix, systemic risk, integrated ownership and control, and the most recent studies on mutual funds holdings structure. The last topic will be discussed in detail with the presentation of new results on mutual funds holdings connections.

Since we are interested in financial markets, we do not delve on theoretical results that are not applied to our specific focus, nor on many applications of complex networks on Economics: such as GDP, considering clustering (Ausloos and Gligor 2008; Gligor and Ausloos 2007, 2008), focusing on the dynamic evolution of the system (Miskiewicz and Ausloos 2006, 2010), and Granger causality (Caraiani 2013), just to cite a few.

For the same reason, we are not extending our review to the rapidly expanding literature on International Trade Network (Bhattacharya et al. 2008; De Benedictis and Tajoli 2011; Garas et al. 2010; Garlaschelli and Loffredo 2004; Schweitzer et al. 2009).

2 Correlation Matrix

How strongly correlated are the stock markets? What is the level of market randomness and dependence? How does the structure change during expansions and recessions? These are the main questions addressed by the papers that use complex networks for the study of the correlation matrix. Correlation matrices play a relevant role in the paramount financial problem of optimal portfolio selection (Elton et al. 2014; Markowitz 1952). In Econometrics, several tools have been developed for their analysis: from ARCH/GARCH models to vector autoregression, principal component analysis and copulas. The perspective of complex networks, besides offering a different approach for a proper correction of the correlation matrix (Aste and Di Matteo 2010; Pantaleo et al. 2011), mainly uses the network approach to build a network structure among financial quantities, and introduces a distance inversely dependent on correlation.

In 1999, Mantegna proposed the distance is $d_{ij} = \frac{1}{2} \sqrt{1 - \rho_{ij}}$, where $\rho_{ij} \forall i, j$ are the correlation coefficient computed between all pairs of stocks of the portfolio by considering the synchronous time evolution of the difference of the logarithm of daily stock price (Mantegna 1999).

It can be proved that d_{ij} is a mathematical distance. Another benefit of using d_{ij} instead of ρ_{ij} is that the most correlated stocks are the closest, which means that the distance is the shortest. Since ρ_{ij} are gathered into a matrix, also d_{ij} constitute the distance matrix D (Caldarelli 2008).

Soon, it was clear that correlation matrices, as well as their deterministic transform into distance matrices, are far from being random networks (Bonanno et al. 2003; Caldarelli 2008). Therefore, it was straightforward to look for the origins of the dependence. Of course, D constitutes a complete network, gathering too much information. A proper analysis must evidence and filter main features and characteristics. One of the most used quantities for such a filtering is the Minimum Spanning Tree (MST). The MST is a sub-network that keeps all the n nodes of the network, but only the $n - 1$ links with the minimum weight, provided that the connected components remain connected. Originally used to detect the lightest routes on a graph, it is calculated through a recursive algorithm. One advantage of using the MST is the possibility of building proper visualization of the structure of the closest stocks in distance matrix D .

The minimal spanning tree (MST) is attractive because it provides an arrangement of stocks, which selects the most relevant connections of each element of the set, and hierarchies can be settled. In Mantegna (1999), the technique is applied to a portfolio of stocks of the S&P 500 index, and it provides a taxonomy that shows the clustering of many groups of stocks, which are homogeneous from an economic point of view.

An interesting aspect on the network analysis of correlation is given by the progressive change of the graph structure, as the time horizon decreases, from a complex organisation to a simple form (where clusters are sparser), so adding further insights and empirical evidence for discussion of the hypotheses of dependence and independence that are most used in financial market models (Bonanno et al. 2004). The MST also proves that during crises the distance among markets decreases (Sandoval and De Paula Franca 2012), and investment signals may be detected (Brookfield et al. 2013). Further techniques for filtering information from the correlation network, like the Planar maximally filtered graph, have been explored to overcome the strong dependence of the presence of links in the MST on the time lag selected for the analysis (Pozzi et al. 2008).

The MST has been applied also to time series of global currencies. In this case, the geographical proximity plays a key role in showing the differences of European and Asian clusters, and the interdependence of the currencies of countries at E.U. borders. As expected, the key currencies belong to major economic countries and the U.S. dollar plays the role of primary currency for its remarkable influence. Therefore, each currency depends on the U.S. dollar and on the key currency of the region where this belongs. The predominance of the U.S. dollar is also proved in other studies (Naylor et al. 2007), which analyse a different sample of worldwide currencies and use another metric distance function, the Gower one (Gower 1986). This result is in line with the U.S. hub role detected by the MST in the international trade networks (Maeng et al. 2012).

The MST has been applied to the distance matrix calculated on correlations among GDP. The structural topology of the MST, sampled at different times, allows the identification of different clusters of countries based on their indebtedness and economic ties. The main results show that with the debt crisis, the less and most affected Eurozone's economies are shaped as a cluster in the MST. In recent papers

(Ausloos and Miskiewicz 2010; Miskiewicz and Ausloos 2010), MST, entropy and other indices are used to prove that the mean distance between the most developed countries, decreased from 1960 to 2000, which can be considered a proof of economic globalization of these countries.

It can be concluded that the correlation matrix has inspired several studies, mostly conducted by physicists, and has contributed to the development of the study of distances, clusters and induced hierarchies on networks.

3 Systemic Risk

Banking has attracted a number of dedicated studies, mainly fostered by the 2008 subprime crisis. Terms like “Too big to fail” soon became part of common talks, and had a relevant role in the public debate on bank saving policies for the recent crises.

In finance, systemic risk is the default risk of an entire financial system. The concept expressed by the phrase “Too big to fail” (TBTF) is that a single financial institute may hold so much credit, that saving it -instead of letting it fail- becomes economically convenient to prevent the failure of the entire financial system. The phrasing was already in use when the crisis in 2008 made the concept prominent and gave a big impulse to the reform of financial legislation (White 2014). TBTF financial institutions are not necessarily banks: in principle, any company that primarily holds financial instruments (such as stocks, bonds, loans, derivatives, etc.) as assets on its balance sheet is exposed to the risk of debtors’ insolvency. The term *cascades* outlines that the insolvency of one institution has negative consequences on others, becoming a *contagion* if the outcome is as bad as causing their insolvency. The larger the financial institution, the worse the effect on the economy, the more likely is the decision of policymakers to intervene by providing support to the financial institutions. The drawback is that these actions also create moral hazard and expectations for the institution’s owners and managers, opening the door to possibly even bigger and deeper crises.

The development of a regulatory system needs a clear understanding of risk exposure and its monitoring. In order to achieve these goals, scientific literature has mainly exploited the concept of centrality on the network, and simulation models of contagions.

Studies on the topology of networks have led to the development of techniques for ranking the *centrality* of nodes in a network. The more a node is *central*, the more it is relevant for the property under observation. Such rankings are often referred to as *centrality measures*, although most of them are not measures in accord to mathematical terms.

Financial networks mirror regional and sectorial organization (Allen and Babus 2009; Bellenzier and Grassi 2013; Bellenzier et al. 2015). Therefore, the empirical estimate of quantities that are standard in complex networks—clustering coefficient and shortest path length besides centrality—has highlighted regional disparities (Boss et al. 2004).

Specific features of banks can be captured by ad hoc models that assess the systemic relevance of a given institution to the contribution of heterogeneity in network structures and concentration of counterparty exposures (Cont et al. 2013; Bougheas and Kirman 2014).

The problem of the optimal network design is quite relevant for the propagation of crises (Leitner 2005; Bougheas and Kirman 2014; López-Pintado 2006), which cause financial earthquakes when triggering responses from a large portion of the financial system (Vitting Andersen et al. 2011). Usually, empirical estimates require large data set that are not so easy to retrieve and manage. Specific network-based measures for ranking the relevance of nodes have been developed (Battiston et al. 2010, 2012; Bellenzier et al. 2015), suggesting that the debate should include issues eventually even more serious than TBTF, such as *Too central to fail* (impacting those who are important via network effects) and *Too correlated to fail* (similar portfolios and/or strategies).

Likewise, in *Too interconnected to fail* (TITF) (Markose et al. 2012) the 2007 credit crisis was empirically reconstructed through data. Dense clustering and mutual exposure identify the TITF institutions, where super spreaders dominate in terms of network centrality and connectivity. Studies focusing on the role of shocks on the overall stability of the financial system (Allen and Gale 2000; D’Errico et al. 2009; Gabbi et al. 2012; Steinbacher et al. 2013) and policies for market regulation are well represented in literature (Gai et al. 2011; Gai and Kapadia 2010; Halaj and Kok 2015).

We can conclude this section with a remark about the above-mentioned models of financial contagion: they do not constitute the only contribution to understanding the extent and consequences of the failure of a financial institution as pinpointed in (Bougheas and Kirman 2014).

4 Integrated Ownership and Control: Complex Networks of the Shareholding Matrix, and Directorate Interlocks

Indirect ownership is quite a relevant issue, mostly when dealing with antitrust measures, and it is very relevant for detecting Chinese boxes and tunnelling. In order to outline the phenomenon, let us consider a company A that does not buy directly shares of a company B (direct ownership); but A holds s_{AC} shares of an intermediary company C that, in turn, owns s_{CB} shares of B. In this way, no direct ownership of A in B is recorded, but, actually, A owns $s_{AB} = s_{AC} \times s_{CB}$ share (indirect ownership). Given the cross-shareholding matrix $A = (s_{ij}) \in R^{n \times n}$, the ownership through one intermediary is given by the matrix product $A^2 = A \times A$. Analogously, the ownership through two intermediaries is given by $A^2 = A \times A$, and, in general, the ownership through n intermediaries is given by A^n .

Therefore, the total ownership, through direct ownership, and through any number of intermediaries, is given by $Y = A + A^2 + \dots + A^n + \dots = (I - A)^{-1}A$,

where I is the identity matrix. Corrections to prevent double counting due to loops (Chapelle and Szafarz 2005, 2007) give rise to the integrated ownership matrix $V = \text{diag}(I - \bar{A})Y$, where $\bar{A} = (\bar{a}_{ij})$, and $\bar{a}_{ij} = \sum_{k=1}^n s_{kj}$, so the value depends only on the column j . The elements of $V = (v_{ij})$ represent the number of shares that the company in column j holds in the company in the row i , counting both direct and integrated ownership through any possible path in the networks. Issues on convergence do not raise because $s_{ij} < 1, \forall i, j = 1, \dots, n$. Results are quite different, depending on the country. Italian companies listed in the MIB30 index do not show long chains of control, and the existing ones can be easily explained following the raise and settlement of the single companies. On the contrary, the Japanese market shows clear signs of tunnelling, and some examples have been detected for the German market (Flath 1992).

The availability of large databases has moved the investigation from small national data set to the international ownership network and techniques from complex networks have produced further results. For instance, the estimate of the assortativity coefficient shows the strong tendency to form high-connected groups (Rotundo and D’Arcangelis 2010a).

Transnational corporations are confirmed to form a giant bow-tie structure, where the central nodes belong to a strongly connected component. Nodes in the strongly connected component are connected by cross-shareholdings, since each of them owns some shares of the others. Such strong component constitutes a small tightly-knit core of financial institutions, eventually TITF; therefore, raising important issues on market contagions, resilience and concentration both for researchers and policy makers (Rotundo and D’Arcangelis 2010a, b, 2014; D’Arcangelis and Rotundo 2014; Rotundo 2011; Bougheas and Kirman 2014; Vitali et al. 2011). For instance, the node out-degree is a quantitative measure of portfolio diversification of the company corresponding to the node. The detection of the eventual power law decay in the histogram of the quantities of interest becomes a standard estimate in Econophysics, since it opens the way to models. Quite interestingly, the histogram of the out-degree and the change of the value of the exponent of the power law clearly show the disappearance of the middle-sized investments through ownership (in favour to the return to the core business), in line with practical managerial issues more than to instances of market expansion.

The concept of integrated ownership is quite different from control. Let us consider the following example: a chain of ownership A-B-C where $s_{AB} = 51\%$ and $s_{BC} = 51\%$. Thus, $s_{AC} = (51)^2 \cong 26\%$. This means that A controls B, B controls C, but A does not have a sufficient number of shares to control C. Therefore, different ways to achieve control must be considered (Chapelle and Szafarz 2005, 2007; Rotundo and D’Arcangelis 2010b).

Interlocked directorates refer to the practice of members of corporate boards of serving on the boards of multiple companies. As early as 1969, the debate on the flaw of fair competition through the interlocking directorate was quite active, especially in reference to laws enacted in 1914. In fact, the inter-organizational élite co-optation can be seen as a cooperative strategy between economic organizations for reducing sources of uncertainty.

Empirical evidences show that links -although dynamically evolving- are persistent, with a stable core, and they involve companies managed by families strong at the local level (Bellenzier and Grassi 2013). A recent analysis (Rotundo and D’Arcangelis 2010b) shows that companies listed in the stock exchange for a long time share the board of directors, whilst newcomers enter the market buying shares. Interlocks among companies can be described as links, and the analysis of the subsequent network is straightforward.

The application of centrality measures shows positive correlation among the rank of interlock and firm value, but positive correlation with betweenness and flow-betweenness, representing the intensity of the relationship between companies, capturing the volume of information flowing from one company to another through the interlocks (Crocì and Grassi 2014; Grassi 2010; Grassi et al. 2008).

Understanding a complex system is quite different from controlling it. We may conclude that the role of social interaction is quite relevant, since the board of directors have their weight in achieving control (Chapelle and Szafarz 2005, 2007; Rotundo and D’Arcangelis 2010b) and further studies are available on the personal connections among important managers.

5 Investment Decisions and Institutional Investors

Complex network methodologies have been applied to the field of financial markets for many purposes. The methodology is suitable to highlight the impact of networks of investors and managers on investment decisions or directly on stock prices. Besides the applications to the interdependence of stock markets, to systemic risks and to integrated ownership and control, complex network methodologies have recently begun to be used in the area of investments and managed portfolios.

Aiming to construct an index of attractiveness of various capital markets, a recent research (Cetorelli and Peristiani 2013) carries out an in-depth analysis of the “patterns of relationships” among financial centres. Using network analysis, the Authors show that although the London Stock Exchange, the Deutsche Börse, and the Hong Kong Stock Exchange became more competitive, the U.S. exchanges remained the favourite destination for foreign issuers which wish to cross-list on multiple exchanges. Along the same line, other Authors (Lucarelli et al. 2012) apply network centrality measures to the indirect network between trading venues (regulated Stock Exchanges and Alternative Trading Venues, ATVs) with the aim to observe the dynamics of simultaneously traded European stocks from 2005 to 2009. Their results show that the advancement of Alternative Trading Venues eroded the isolated-centrality of major Stock Exchanges (above all London); in contrast, degree-centrality significantly increased for the majority of the Stock Exchanges analysed in their sample. The introduction of multi-trading venues does not deteriorate connectivity of cross-listed relationships, unveiling that multi-trading co-exists with secondary market cross-listing.

The investment management industry has been playing an increasingly important role in the financial system, especially in the most advanced economies. In recent decades, credit intermediation has been progressively shifting from the banking to the non-banking sector, particularly to the asset management industry (through different investment vehicles such as mutual funds, hedge funds, exchange-traded funds, private equity funds, pension funds). Focusing on mutual funds, the number of mutual funds in the U.S. reached 79,669 in 2014 with a value of 31.38 trillion U.S. dollars for the total assets under management.¹ U.S. mutual funds account for roughly half the global asset under management in the world. Europe has an equally important role with 9.576 trillion of asset under management. Such a crucial role for the asset management industry has obvious paybacks for financial intermediation: investors gain a better diversification of their portfolios and grant a more stable financing of the real economy even during periods of distressed market conditions. Other benefits of the asset management vehicles over banks concern the stability of the financial system: the banks are exposed to solvency and liquidity risks (due to their typical short term funding) whereas the investment risk of the shares issued by mutual funds relapses on end investors.

Focusing on this field, complex network methodologies have been firstly used to analyse how information is disseminated among mutual fund managers in financial markets and how the diffusion of such news influences stock prices. Focusing on connections between mutual fund managers and corporate board members via shared education (the connection is their attendance of the same school), a research (Cohen et al. 2008) identifies the transmission of insider information and finds that managers favour the investment on companies they are connected through their network. Placing larger bets on “connected firms”, which have performed significantly better than non-connected ones, mutual fund managers have significantly improved their performance.

The availability of data on holdings of a sample of US actively managed equity mutual funds (Augustiani et al. 2015) allows other Authors to examine the effect of mutual fund connections, through managerial sharing, on performance and stock holding commonalities. Their analysis of return correlations and portfolio holdings shows that more interconnected funds managers tend to buy and sell similar stocks, hence increasing the similarity of portfolio holdings and undermining the distinctiveness of their investment strategy. Assessing performance effects, the Authors find that highly connected funds significantly underperform weakly connected funds by about 1.4 % on a yearly risk-adjusted basis. Conversely, fund family performance remains almost unaffected by the intensity of fund connections, and greater fund connections can significantly enhance family-level profit margins.

A subsequent research focused on tie between analysts and companies (Cohen et al. 2010) investigate the dissemination of information in security markets through the recommendations of sell-side equity analysts used to study the impact of

¹The largest fund management companies worldwide as of December 2014 are Blackrock, Vanguard Asset Management, State Street Global Advisors and Fidelity Investments.

social networks on agents' ability to gather superior information about firms. The hypothesis is here that school ties between analysts and senior corporate officers provide comparative information advantages in the production of analyst research. The main result is that equity analysts outperform on their stock recommendations when they have an educational link to that company. Results are strong, so much that a strategy of going long (for the recommendations to purchase given by analysts with school ties) and going short (for the buy recommendations of analysts without ties) returns a performance of 6.60% per year. The Authors' conclusion is that analysts' social networks facilitate the direct transfer of information, or alternatively that these networks simply allow analysts to better assess managerial quality. Even if the Authors do not always use complex network measures, the merit of these papers is to highlight the importance of network structures and interactions among agents in the analysis of information flow and price evolution in security markets.

The metrics from network analysis allows to analyse the impact of Sovereign Wealth Fund (SWF) equity investments on target firm operating performance (Del Giudice et al. 2014). The aim of the paper is to investigate whether target firms, which are better connected to each other by means of the SWF investments, gain benefits in terms of higher performance. The results indicate that more central firms in the SWF-target firm network have better operating performance and that the effect is related to the size of the stake acquired, is larger if the investment is direct and in the domestic country and if the SWF is run by a politician.

The analysis of the relationship between the location of the fund in its network and the investment performance, risk taking, and flows is the object of a paper focused on pension funds (Rossi et al. 2015). Using data on a large set of UK pension fund accounts over the period 1984–2004, the Authors investigate whether network centrality explains managers' investment performance, risk-taking behaviour, and flows. The centrality of the management company providing the fund is derived from the number of connections it has with other management companies through their commonality in managing for the same fund sponsors or through the same fund consultants. In detail, individual pension fund accounts can be connected by their sharing of the same consultant and/or the same manager. The results show that a fund-manager's (relative) degree of centrality in a network positively affects risk-adjusted returns and growth in assets under management and that this effect is particularly strong for large fund managers, even after controlling for size. Once a central position has been established, the manager tends to reduce risk-taking behaviour and reduces the chances of getting fired by institutional clients.

The fact that many mutual funds around the world have suffered from negative returns during the global financial crisis emphasizes how volatility can rapidly spread among previously unrelated assets in times of high uncertainty and turbulence. This observation provides the starting point for an investigation (Azmi and Smith 2010) of the spread of the current crisis in the correlation networks amongst a sample of mutual funds across seven regions globally. Using the data of equity funds in ten countries representing seven regions, the Authors select two funds from each country based on the highest net asset value and built a correlation network of the weekly mutual fund log-returns over the period from April 7, 2006 to April 27,

2009. They show that the losses in financial assets within certain countries could spread and follow a cascade or epidemic flow like model along their correlations. A first conclusion that can be drawn from the analysis is the rapid spread of the credit crisis amongst previously uncorrelated markets and countries. The correlation networks under examination do not cause the transmission chain of collapse, but they are tied to it. Such architecture encourages the excesses of the global financial crisis, motivates aggressive risk taking and pushes some asset prices to unsustainable levels, increasing financial fragility.

A second stream of research uses complex networks techniques in order to detect connections among mutual fund holdings and relates complex networks measures to the dynamic of risk and return. The collection of the ten largest positions of 18 Vanguard and Fidelity family funds (Solis 2009) provides a new approach to visualizing the way stocks are affiliated to mutual funds as a bipartite graph, and computes network summary statistics. The stock network has a high clustering coefficient (indicating “prominent” stock hubs), which suggests that the managers’ selection of stocks is not made independently as if the network were that of a purely random graph with similar expected number of links. The higher diameter and average degree distance between two vertices (6 and 2.91 vs. 3 and 2.01 for the random graph) suggests a small-world behaviour, due to the highly connected network of stocks, mostly blue-chips, which populate the mutual funds sample.

In a more complete study, focusing their attention on the indirect connections among holdings resulting from common ownership of a sample of mutual funds from 1980 to 2008, Anton and Polk (2013) find that pairs of stocks held in many mutual funds’ portfolios show future excess correlation between stock returns. The Authors demonstrate that this “common ownership effect” is stronger for common owners who are experiencing extreme positive or negative flows in low-float stocks. Based on these results, the paper supports a cross-stock-reversal trading strategy that exploits the information in ownership connections and generates significant abnormal returns of more than 9% per year, controlling for market, size, value, momentum, and other characteristics. Following a similar approach in identifying pairs of mutual funds linked by common portfolio holdings, Blocher shows that spillover effects associated with the fund flows of an investor’s network neighbours account for roughly 2% per quarter and are the result of crowded trades, since they are completely reversed in the subsequent year (Blocher 2013).

In the same stream, Braverman and Minca demonstrate that the network of common asset holdings is useful to identify systemic funds (Braverman and Minca 2014). Using quarterly equity mutual fund holdings data ranging from January 2003 to December 2012, they analyse the interrelations due to common asset holdings and construct a measure of fund vulnerability for the shocks of their neighbours in the network. The Authors demonstrate that this “vulnerability index” is useful in predicting returns in periods of mass liquidations, since it helps in identifying vulnerable funds based on asset holdings and the liquidity characteristics of the stocks.

Guo, Minca and Wang analyze the topology of the network of common asset holdings, a network in which nodes represent managed portfolios and edge weights

capture the impact of liquidations (Guo et al. 2015). Focusing only on the sub-graph of weak links (those that lead to significant liquidations), the Authors analyze the degree centrality and find that this measure follows power law distribution and is correlated with returns. For individual portfolios, higher degree is associated with future higher return in the long run, but it may negatively affect the performance during financial crises. At the aggregate level, stronger connectivity among portfolios is associated with higher systemic risk. Exploring network clustering, they identify a small number of communities, densely linked, that concentrate a significant proportion of the portfolios.

D’Arcangelis and Rotundo explore the commonalities in the holdings of Italian funds investing in domestic stocks (D’Arcangelis and Rotundo 2014). Following the empirical evidence that shows that fund managers take common decisions on stock holdings, both for benchmark constraints and for style management decisions, the Authors use the methodology of complex network analysis, to describe the way in which stocks are related to mutual funds and to detect the implications of the interactions. The results highlight a large core group of portfolios that have many stocks in common; while other funds invest in a wider variety of stocks. These results confirm empirical findings on US market (Solis 2009) and show that Italian mutual funds holdings are highly interconnected, suggesting the existence of a small-world behaviour and the tendency of mutual funds managers to steadily invest in a restricted number of well-established high capitalization stocks (blue chips). They also test the impact of overlap on performance and risk (raw performance, Sharpe Ratio, standard deviation, beta coefficient and fund tracking error), and find that the overlap is predictive of performance similarity, at least for the samples involving blue chip stocks, even though comparison with and influence of many other factors are not to be completely disregarded. The intersection of small cap holdings, usually exploited for tactical asset allocation purposes, is not feasible to conveniently differentiate final performance of funds. The results support the thesis of substantial passive management of institutional portfolios, realized through the investment in a portfolio of a limited number of stocks, which are selected by mathematical algorithms in order to optimize the solution of the index-tracking problem.

5.1 Mutual Funds Holdings: A Network Analysis

The data used for the analysis are the holdings of 215 European mutual funds, as listed in the database made available by Morningstar Italy in December 31st, 2015. The funds legal address is in Europe, and their investment focus is the European equity market. In detail, the parameters used in the data query from Morningstar request that the funds belong to one of the following categories: EURO Large Cap, Europe Large Cap Blend, Europe Large Cap Growth, Europe Large Cap Value. The Bloomberg database has been used to check the equity exposure for each of the 272 mutual funds coming out from the query. This check caused the reduction from 272

Table 1 Funds primary benchmark of the funds of the sample (Source: Bloomberg)

Fund primary benchmark	No of funds	Sample (%)
MSCI Europe NR USD	96	44.65
MSCI EMU NR USD	72	33.49
MSCI Europe Value NR USD	29	13.49
MSCI Europe Growth NR USD	17	7.91
MSCI Europe NR EUR	1	0.47

to the final 215 list of funds, because some funds do not authorize the disclosure of the composition of their portfolios. The selected funds belong to investment houses of 14 European countries.²

For each fund in the different categories, we have registered all the stock holdings ranked in terms of weight: the equity sample contains 1603 stocks belonging to 51 different countries.³ The sampled funds adopt similar benchmarks, confirming their belonging to comparable categories: 92.5 % of funds are related to MSCI Europe Indices (44.2 % to MSCI Europe NR and 21.4 % to MSCI Euro NR EUR), STOXX Europe Indices and FTSE Indices, 4.2 % are not benchmarked and only 3.26 % are related to minor benchmarks. Following the benchmark attribution by the site Morningstar Italy, 44.65 % and 33.49 % of the funds are represented by the MSCI Europe and by the MSCI EMU indices (Table 1).

We have started our analysis by building a network on the mutual funds stock holdings. A link is drawn from mutual fund i to stock j if mutual fund i owns stocks of company j . Links are only drawn from a mutual fund to a stock. Therefore, starting and ending nodes of each link belong to different sets. Networks showing this property are named bipartite.

We represent the funds-stocks network through a matrix A . Rows of $A = (a_{ij})$ correspond to funds, and columns relate to stocks. The weights of the network links a_{ij} are the percentages of stock j held by fund i . In our sample, matrix A has 215 rows and 1603 columns. The binary matrix $B = \text{sign}(A)$ is well suitable for projecting the network into the space of relationships among mutual funds as follows: two funds are connected if they have at least one stock in common. The funds-funds matrix C is then created through matrix multiplication: $C = BB^T$. For any two funds i and j , the elements c_{ij} report the number of stocks common to both funds i and j .

Lastly, the stocks-stocks matrix $D = (d_{ij})$ is defined as $D = B^T B$. For any two stocks i and j , the elements d_{ij} report the number of funds that own shares of both stocks i and j .

²Five European countries mostly contribute to the group of 215 mutual funds: Germany (18 funds), France and Italy (45 funds each), United Kingdom (59 funds), Switzerland (26 funds). Other 22 funds belong to investment houses of 9 other European countries. Due to the exiguity of the samples, they have been gathered into the category "Others".

³In the sample of stocks, the ones belonging to the five most represented countries (DE, FR, IT, SW, UK) account for the 57.08 % of the sample.

5.1.1 Analysis of the Overall Funds-Funds Network

The funds-funds network is a non directed matrix, as witnessed by the symmetry of matrix C , which shows the overlap of investments among fund i and fund j . However, this information is biased by the dimension of both funds i and j , since funds with more holdings could be more overlapped only due to the number of different stocks in which they are invested. Considering matrix A , the size of funds can be calculated at once as $(\text{sign}(A^T) e)$, where $e = (1, 1, 1, \dots, 1)^T$ is a vector with 1063 components. In matrix B , the size of each fund is reported in the elements on the diagonal. In order to overcome the issue, we have divided each element in matrix C by the maximum between the dimension of funds i and j , so obtaining the symmetric matrix $F = c_{ij}/\max(b_{ii}, b_{jj})$.

The elements on the diagonal were set equal to zero to avoid counting the size of the fund. Each of the 215 rows of F represents the average overlap of the fund on row i with the remaining 214 funds of the sample. The result is a vector of 215 cells (each representing the mean overlap of fund in a row with the other funds in the columns), whose statistics (mean, median and mode, skewness and kurtosis) are summarized in Table 2 (column full sample).

The analysis has been repeated calculating the mean of each of the rows of F within the columns of funds belonging to the single country (DE, FR, IT, SW, UK and “Other”); this average represents the mean overlap of the fund of row i with the other funds of that country. The statistics on these country vectors are shown in the respective columns of Table 2.

The Jarque-Bera test⁴ is used to check the null hypothesis that data come from a normally distributed population. The results show that normality holds on the full sample and does not hold on subsamples.

Table 2 Mean overlap of the F normalized funds-funds matrix

Funds-funds matrix (normalized sample, average values)							
	Full sample	DE	SW	FR	IT	UK	Other
Mean	0.0526	0.0643	0.0577	0.0537	0.0498	0.0499	0.0475
Median	0.0545	0.0656	0.0608	0.0553	0.0497	0.051	0.0605
Mode	0.0606	0.0397	0.0514	0.0514	0.0514	0.027	0.0606
Skewness	-0.0282	-0.007	-0.3082	-0.188	0.3507	0.1348	-0.4827
Kurtosis	2.0354	1.9451	1.9045	2.0763	2.5952	1.7755	1.8971
Jarque Bera							
h	1	0	0	0	0	1	1

⁴If $h = 1$, the hypothesis of normality is rejected, $h = 0$ means that it is accepted.

5.1.2 The Analysis of Centrality

On the basis of this sample, we have calculated nowadays standard measures of centrality on networks. At a first analysis, the node degree or “degree centrality” (that indicates the number of neighbours of a node) reveals that our network is nearly a complete one. The betweenness centrality is the number of shortest paths going through the node under examination in relation to the total number of shortest paths of the network. As this measure of centrality indicates if a node can be “intermediary” within the network, it assumes a specific meaning for portfolios like mutual funds. A high value of the betweenness centrality unveils funds that “bridge” different groups: a fund with the higher betweenness owns stocks that are owned by different groups, which would not overlap otherwise.

Whether the network is not showing well distinct groups, the fund with the maximal betweenness is not investing in the core investment target of most mutual funds, so its eventual crash is not as relevant in other applied problems, for instance in the spread of epidemics or in the true collapse of a bridge. This is exactly our case. The entire group is so highly connected that the values of the betweenness centrality do not provide an informative discriminant analysis. Similarly, those funds high on eigenvector centrality are linked to well-connected portfolios and may influence many others in the network either directly or indirectly through their connections.

Another centrality measure that can be applied to the funds-funds matrix is the k-shell decomposition (Garas et al. 2012), which tries to detect the most influential nodes in the network of funds. Therefore, our analyses on K-shell does not lead to meaningful partitions. The application of the Louvain method of community detection (Blondel et al. 2008) reaches a similar conclusion.

In conclusion, results state that the funds in our network are highly connected and their portfolios exhibit close overlaps. These results can be the consequence of the adoption of similar investment policies in presence of different but highly correlated benchmarks.

5.1.3 Analyses Considering the Benchmark

Such a conclusion suggested we should look into it further. Therefore, we have repeated the analysis on two sub-portfolios obtained by insulating the stocks belonging to the benchmark and those out of the benchmark. The aim is to detect the possibility of different behaviour of fund managers in the management of the market based and the tactical components of the portfolios under management.

We have extracted the constituents of the MSCI Europe index from the MSCI site⁵: of the 442 constituents of the MSCI Europe, as many as 440 of them are present in the sample of 1603 stocks of our funds-stock matrix. Then we have built two separate networks: the funds-stock in the benchmark (215 rows \times 440 columns)

⁵ Available at <https://www.msci.com/constituents>.

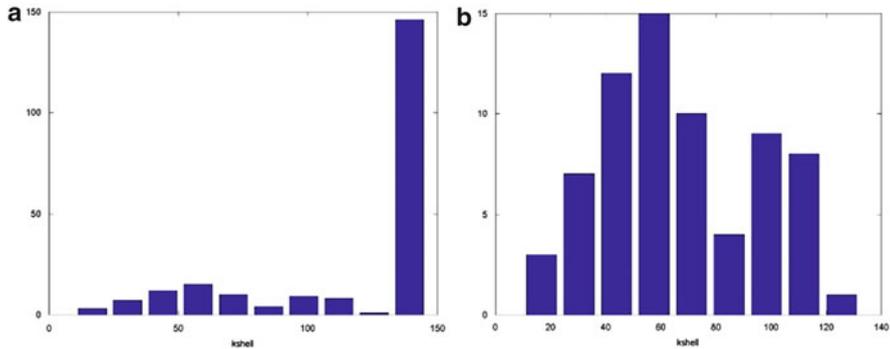


Fig. 1 Bar diagram of the k-shell of the funds-funds network calculated on the out-of-benchmark stock sample. (a) Entire distribution (b) having removed the peak at 146

and the funds stocks out of benchmark (215 rows \times 1163). Looking at the results on the first subsample, the conclusions on the degree centrality and on the k-shell decomposition do not change much. The fluctuations of the values do not lead to significant differences. A quite different picture emerges when considering the second network built on out-of-benchmark stocks. The k-shell degree shows a large group of 146 funds showing high peak at 146, a second group of 18 funds with values between 122 and 100 and a residual group of funds with k-shell under the value of 100 (see Fig. 1a, b).

This is quite different from the nearly complete connection of the previous networks, and also quite far from the ubiquitous power law behaviour. In fact, there are many nodes with a high node-degree. This means that the connections among funds are very dense also in the out-of-benchmark sample, which means that even if the investment is diversified with a tactical and unsystematic component of the portfolio, such diversification is not strong enough: a situation completely different from what expected in presence of a power law distribution (for an example, refer to Fig. 2b of D’Arcangelis and Rotundo 2014).

Starting from the percentage of weights of the five most significant countries of the sample,⁶ a geographical analysis of the three k-shell groups reveals that

- apart from Great Britain funds, whose frequency falls from 27.4% in the full sample to 21.9%, the subsample of the most connected stocks (k-shell value equal to 146) does not signal any outlier;
- the percentage incidence of UK holdings increases with the decreasing value of k-shell, supporting the existence of much *bigger diversification* benefits in UK funds;

⁶Germany 8.4%, Switzerland 12.1%, Italy and France 20.9% each, UK 27.4% and the residual Countries 10.1%.

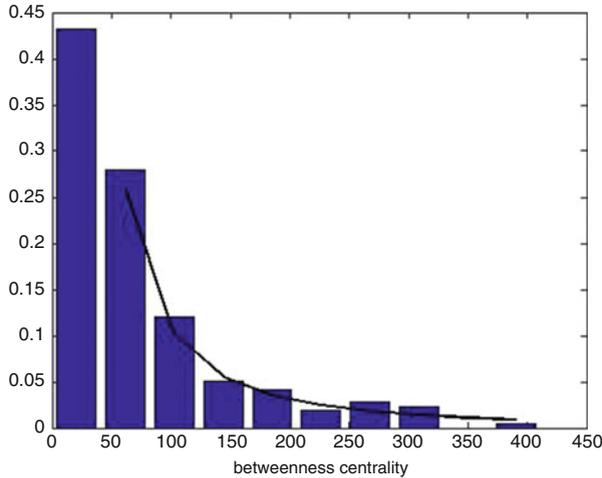


Fig. 2 Histogram of the betweenness (fund-funds matrix on out-of-benchmark sample). The tail of the distribution is well fit by a power law $f(x) = cx^{-\alpha}$ with coefficients (with 95 % confidence bounds): $c = 488.7$ (27.2, 950.1), $\alpha = 1.81$ (1.589, 2.031). Goodness of fit: SSE: 0.0005668, R-square: 0.991, Adjusted R-square: 0.9897, RMSE: 0.008999

- Italy and Germany show an antithetic behaviour relative to UK, with a strong underweight in the group of funds with minimum k-shell (under 100), supporting the evidence of more concentrated portfolios.

In contrast, the betweenness centrality shows a power law behaviour (see Fig. 2). This means that there is a continuum of values in the relative overlap of funds, with the usual implications when it comes to resilience to spread of volatility, contagions and financial fragility.

5.1.4 A Country-Based Analysis of Fund-Fund Network

Moreover, the analysis of the funds-funds matrix was performed also on separate subsamples of funds belonging to each of the five single countries under examination (Germany, Switzerland, France, Italy and UK, and the residual class “Others”). In this case, indeed, the mean has been calculated for each row of the funds-funds matrix, whose dimension is related to the number of funds of the country in the network. The result of this operation is a vector with 215 components, where each value shows the mean of the stocks that overlap the fund in the rest of the sample (the elements on the diagonal were set equal to 0). Table 3 shows the mean of this vector.

Germany funds qualify as the most connected network. In fact, the German funds in the subsample funds-funds matrix (18 rows and columns) have the highest overlap (0.0781 is the highest value on the main diagonal, and also the maximum

Table 3 Mean overlap on the single country normalized funds-funds matrices

Country	DE	SW	FR	IT	UK	Others
DE	0.078	0.068	0.069	0.057	0.061	0.066
SW	0.068	0.063	0.062	0.055	0.055	0.051
FR	0.069	0.062	0.057	0.049	0.049	0.049
IT	0.057	0.055	0.049	0.047	0.050	0.046
UK	0.061	0.055	0.049	0.050	0.047	0.046
Others	0.063	0.048	0.047	0.044	0.044	0.051

Table 4 Mean overlap of the single country stocks-stocks matrices

	DE	SW	FR	IT	UK
DE	3.052	1.978	2.954	1.442	1.497
SW	1.978	2.291	1.930	0.961	1.644
FR	2.954	1.930	3.223	1.423	1.441
IT	1.442	0.961	1.423	0.944	0.717
UK	1.497	1.644	1.441	0.717	1.293

element in the matrix). Moreover, the overall connection of German funds to the funds of the other countries is the higher (the sum of the connections to FR, IT, SW, UK and “Others” is equal to 0.3211, and is highest than all the other sums). Similar conclusions hold for the Swiss funds, with the second highest value on the main diagonal, that shows the total connection among the Swiss funds. The overall connection to the funds of other countries is the higher and equals 0.29. Italian funds reveal an opposite behaviour, showing the lower value of internal connection (0.443) and also the minimum overlap with the funds of the other countries (the sum of the connections with German, French, Swiss, UK and “Other” funds equals 0.253).

5.1.5 Analysis of the Overall Stocks: Stocks Network

The stocks-stocks matrix D is defined as $D = B^T B$. For any two stocks i and j , the elements d_{ij} report the number of funds that own shares of both stocks i and j . The stocks-stocks network is a non directed matrix, as witnessed by the symmetry of matrix D .

The mean of each of the rows of D represents the mean overlap of the stock on row i with the remaining of the stocks of the sample, divided by countries. It is the mean of all the values in the submatrix corresponding to the rows and columns listed for each block in Table 4, that sums up the findings of the calculation of the means on the country blocks of the vector of the means of the overlap values in each row.

The results show that German stocks reveal the maximum overlap both with foreign and domestic stocks. The conclusion is that in the sample under examination German stocks are diffused to the utmost degree. A similar behaviour is detectable also for French stocks, which reveal the maximum overlap with other French stocks (France/France = 3.22); the mean value of the overlap with stocks of other countries is 7.748 and is the second value (France has the minimum overlap 0.944 in the

Table 5 Mean overlap on the single country stocks-stocks matrices

	DE	SW	FR	IT	UK
DE	0.915	0.670	0.884	0.670	0.625
SW	1.491	0.801	1.040	1.102	0.722
FR	1.544	1.166	1.736	0.694	1.079
IT	1.082	0.779	1.199	0.942	0.765
UK	0.592	0.500	0.598	0.418	0.432

sample with Italian stocks). Going through the rows of Table 4, we can notice that also Swiss stocks are more connected with the other Swiss stocks (Switzerland-Switzerland = 2.29) than with the stocks of other countries, but the overlap with German stocks is also very high (1.97).

Of the five countries examined, only Italy and the United Kingdom have a value on the main diagonal that is not the maximum in their row; therefore, for these countries, the overlap among domestic stocks is lower than with those of other countries. Italy and United Kingdom stocks are mainly bought jointly with stocks of other geographical areas: the maximum overlap of United Kingdom stocks is with Swiss stocks; the maximum overlap of Italian stocks is with German and French ones.

Lastly, the overlap between Italy and the UK is always very weak. Further, we may add the comment that Italy has a marginal weight in international benchmarks and that the result of the United Kingdom could be partially due to its disposition to minimize exchange rate risks and avoid costly hedging strategies.

Following the methodology used for the funds-funds network, we have built the stocks-stocks matrix on separate subsamples based on the domicile of the funds in the sample (Germany, France, Italy, Switzerland, and UK, and the residual class "Others").

The difference between the country-based analyses in Tables 4 and 5 lies in the different data of each cell in the stocks-stock matrix: in Table 4, the cell contains the mean overlap of domestic and international funds that share the couple of stocks in the corresponding row and column; indeed, in Table 5, such funds are only domestic (ref. to the row). For instance, in the cell corresponding to DE (row) and UK (column) the value 0.625 is the overlap between DE stocks and UK stocks, considered as a block, owned exclusively by DE funds. The results converge towards the conclusions detected from previous tests: French mutual fund managers confirm to be inclined to overweight the percentage of French stocks in their portfolios and are at the same time exposed to German stocks; Italian funds combine mainly French and German stocks, the same conclusion is valid for Swiss managers. The exposition of UK funds to other countries holdings is systematically low (the row corresponding to United Kingdom has the minimum values of the entire matrix), the consequence is that we detect an intrinsic tendency of these actors towards a strong diversification.

The node-degree in the stocks-stocks network is the number of links connected to the nodes. It can be calculated through the $K_{out} = \text{sign}(D) e$, where $e = (1, 1, 1, \dots, 1)^T$

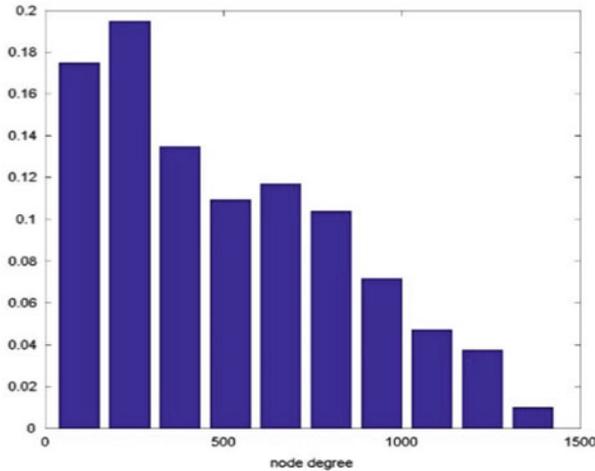


Fig. 3 Node degree of the single country stocks-stocks matrices

is a vector with 215 rows. A node i showing a high degree signals a stock that is associated to many other stocks in the portfolios of the sample mutual funds. A high mean node-degree means that stock i is connected to many others because they appear together in the portfolios of many funds. Figure 3 shows a behaviour different from the power law, although showing a small number of couples of stocks belonging to many funds, and a higher number of couples of stocks belonging to a few funds.

Due to the high connectivity of the matrix, the results of betweenness and k-shells are not performing a clear discriminant analysis; therefore, we have proceeded with the analysis of two well distinct sub-groups: the stocks belonging to the benchmark and the ones out-of-the-benchmark.

5.1.6 Analyses of the Stocks-Stocks Matrix for Benchmark Constituents

On the benchmark, the node-degree is far from the classic power law. Counting the nodes whose degree equals zero, we can notice that there are 1161 stocks that do not belong to the ownership of the same mutual fund with other stocks in the benchmark. The node degree has a peak concentrated on the mode at 428 (343 nodes). Since the remaining nodes have values ranging nearly uniformly from 50 to 437, we may conclude that the stocks in the benchmark either do not belong to the same mutual funds of other stocks or are the most frequently bought with other 343 stocks of the benchmark. Non zero elements in the betweenness show a power law behaviour, which implies that there are a few elements (with high betweenness) that lay in paths connecting communities (Fig. 4).

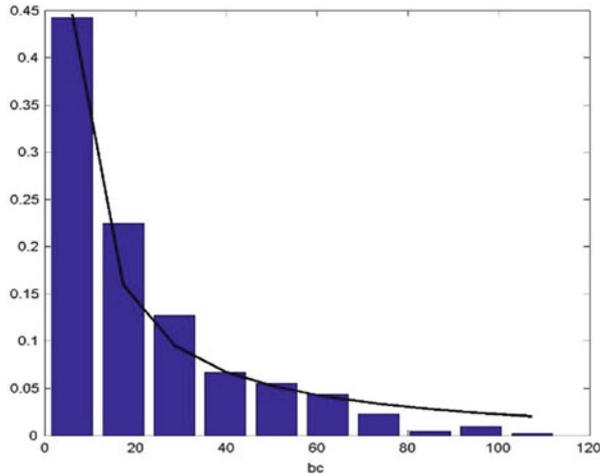


Fig. 4 Histogram of the betweenness for stocks belonging to the benchmark. General model Power1: $f(x) = cx^{-\alpha}$ Coefficients (with 95 % confidence bounds): $c = 2.463$ (1.402, 3.525), $\alpha = 0.9394$ (0.7462, 1.133). Goodness of fit: SSE: 0.0068, R-square: 0.9605, Adjusted R-square: 0.9556, RMSE: 0.02916

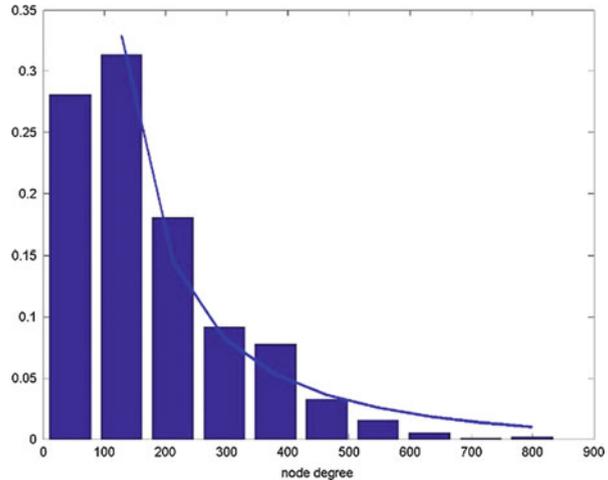
Table 6 Geographical allocation of the community structure—benchmark sample

Group	n.	DE (%)	SW (%)	FR (%)	IT (%)	UK (%)	Other (%)
1	147	17.69	0.68	27.89	11.56	1.36	40.82
2	120	0.00	21.67	0.00	0.00	46.67	31.67
3	88	1.14	17.05	1.14	0.00	50.00	30.68
4	85	28.24	0.00	36.47	8.24	2.35	24.71

The analysis of k-shells shows that there are two very different groups: one at 400 and one at 50. The one around 400 contains 419 elements, confirming the existence of a large group of highly connected nodes.

The presence of groups is well outlined by the Louvain method for communities’ detection (Blondel et al. 2008). Table 6 shows four communities with a specific geographic concentration. Starting from the geographical percentage coverage of the sample of stocks (DE 11.59 %, SW 9.55 %, FR 16.59 %, IT 5.45 and UK 23.64), the analysis shows that Group 2 and Group 3 have a strong concentration on Swiss and UK stocks, whereas the stocks of the other countries (Germany, France and Italy) are concentrated in Group 1 and Group 4. The fact that the UK and Switzerland are non-Eurozone countries may be a tentative explanation for the results that give a snapshot of a sort of “geographical organization” of the communities. The output of the Louvain method for the out-of-benchmark matrix is strongly different, as it does not show a clear geographical allocation within the groups.

Fig. 5 Histogram of the node degree of the stocks-stocks matrix of the out-of-benchmark stocks. General model Power1: $f(x) = cx^{-\alpha}$. Coefficients (with 95 % confidence bounds): $a = 624.9$ (509.7, 1760), $\alpha = 1.559$ (1.2, 1.918). Goodness of fit: SSE: 0.003182, R-square: 0.9643, Adjusted R-square: 0.9592, RMSE: 0.02132



5.1.7 Analysis of the Stocks-Stocks Matrix: Out-of-Benchmark Stocks

The statistics of the stocks out-of-benchmark are quite different from the ones in the benchmark. The node degree has a power law tail distribution. This is in accord with the strategic asset allocation of non-benchmark stocks. In fact, there are a few that are bought by the many mutual funds, and many that are bought by a few mutual funds (Fig. 5). The comparison with the node degree of the entire stocks-stocks network shows lower values in the middle part of the distribution.

The values of betweenness are all concentrated around the mode, with a very few elements with high betweenness. This is in favour of the hypothesis of the presence of groups, with some elements bridging them. The analysis of k-shells shows a quite interesting behaviour, where the highest value at 200 reveals a k-core of overlapping stocks, while the minimum of the k-shell is 3 and the highest number of stocks are in the first bin of the bar diagram (Fig. 6).

The presence of scaling and the absence of well-insulated groups of specific dimension is confirmed also by the Louvain method for communities detection (Blondel et al. 2008). Twenty groups are detected, and the presence of many small groups and a few large ones is clear, although the peaks around 140 and 160 deviate from the power law (Fig. 7). Such groups do not reflect any geographical grouping homogeneity. Therefore, the result is quite different from the result on the benchmark. The absence of country-based grouping confirms the fact that out-of-benchmark stocks are bought by mutual funds mostly for tactical asset allocation purposes without a particular interest in geographical distribution. We leave the investigation of other causes for such grouping to future work.

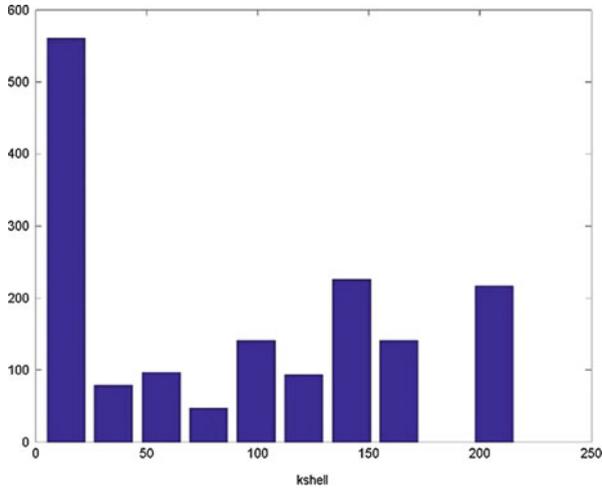


Fig. 6 Bar diagram of the k-shell of the stocks-stocks matrix calculated on the not-benchmark

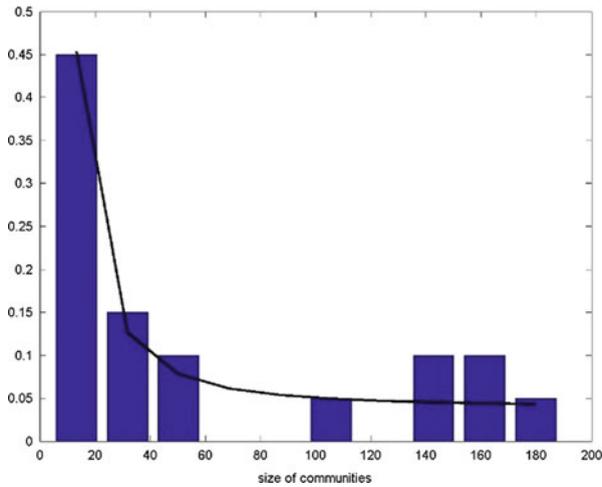


Fig. 7 Size of communities (20 communities with sizes 189, 157, 155, 147, 134, 101, 58, 50, 40, 33, 32, 14, 10, 9, 8, 8, 6, 4, 4, 4). The peaks around 140 and 160 deviate from the power law, although the presence of many small groups and a few large communities is clear. The regression line is $f(x) = b \cdot x^{-\alpha}$, Coefficients (with 95 % confidence bounds): $b = 41.27$ (-138.9, 221.4), $\alpha = 1.781$ (0.06896, 3.494), $c = 0.03919$ (-0.02474, 0.1031). Goodness of fit: SSE: 0.01607, R-square: 0.8996, Adjusted R-square: 0.8709, RMSE: 0.04791

6 Conclusions

The present work has examined a sample of equity mutual funds investing in European stocks and presents various analyses mainly based on the complex network approach applied to stock holdings. The main results show that stocks are connected through the mutual fund owners they have in common and that there are substantial differences in geographical allocation among the different European fund managers. Five larger European countries dominate the market of mutual funds. United Kingdom and Switzerland are the less overlapped, while the highest overlap is among Germany and France. Italy is quite close to Germany and France for the overlap and selection of stocks, but it shows a lower diversification, like the United Kingdom, by looking at funds as a whole. The belonging of UK and Swiss opt-outs of the Eurozone, whose consequence is the exchange rate risk for non domestic investors, could be a probable explanation for our results on community detection that give a snapshot of a sort of “geographical organization” of the core of mutual fund portfolios, the part associated to the benchmark. The results of the different analyses provide valuable input for further research. In particular, we leave to future work the analysis of managers’ behaviour non compliant with geographical allocation of holdings.

Acknowledgements The authors thank COST Action IS1104 for fruitful networking and financial support and Dr. Anna Romagnuolo for editing and proofreading.

References

- Albert R, Barabasi AL (2002) Statistical mechanics of complex networks. *Rev Mod Phys* 74:47–97
- Allen F, Babus A (2009) Networks in finance. In: Kleindorfer P, Wind J (eds) *The network challenge*. Wharton School Publishing, pp 367–382
- Allen F, Gale D (2000) Financial contagion. *J Polit Econ* 108:1–33
- Anton M, Polk C (2013) Connected stocks. *J Finance* LXIX(3):1099–1127
- Aste T, Di Matteo T (2010) Introduction to complex and econophysics systems: a navigation map. In: *Complex physical, biophysical and econophysical systems*, pp 1–35
- Augustiani C, Casavecchia L, Gray J (2015) Managerial sharing, mutual fund connections, and performance. *Int Rev Finance* 15:427–455
- Ausloos M, Gligor M (2008) Cluster expansion method for evolving weighted networks having vector-like nodes. *Acta Phys Pol A* 114:491–499
- Ausloos M, Miskiewicz J (2010) Entropy correlation distance method applied to study correlations between the gross domestic product of rich countries. *Int J Bifurcat Chaos* 20:381–389
- Azmi RA, Smith R (2010) Contagion within financial markets and networks across the globe: evidence from equity mutual funds during the current crisis. Available at SSRN <http://ssrn.com/abstract=1691157> or <http://dx.doi.org/10.2139/ssrn.1691157>
- Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A (2004) The architecture of complex weighted networks. *Proc Natl Acad Sci U S A* 101:3747–3752
- Battiston S, Glattfelder JB, Garlaschelli D, Lillo F, Caldarelli G (2010) The structure of financial networks. In: Estrada E, Fox M, Higham D, Oppo G-L (eds) *Network science: complexity in nature and technology*. Springer, London, pp 131–163

- Battiston S, Puliga M, Kaushik R, Tasca V, Caldarelli V (2012) DebtRank: too central to fail? Financial networks, the FED and systemic risk, scientific reports 2, article number: 541 doi:[10.1038/srep00541](https://doi.org/10.1038/srep00541)
- Bellenzier L, Grassi R (2013) Interlocking directorates in Italy: persistent links in network dynamics. *J Econ Interact Coord* 9:183–202
- Bellenzier L, Vitting Andersen J, Rotundo G (2015) Contagion in the world's stock exchanges seen as set of coupled oscillators. *Economic Modelling* (in press)
- Bhattacharya K, Mukherjee G, Saramaki J, Kaski V, Manna SS (2008) The International Trade Network: weighted network analysis and modelling. *J Stat Mech*, P02002
- Blocher J (2013) The externalities of crowded trades. Technical report, Vanderbilt Owen Graduate School of Management Research Paper No. 1968488. Available at SSRN: <http://ssrn.com/abstract=1968488> or <http://dx.doi.org/10.2139/ssrn.1968488>
- Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) [Fast unfolding of communities in large networks](#). *J Stat Mech Theory Exp* (10), 10008
- Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU (2006) Complex networks: Structure and dynamics. *Phys Rep* 424(4–5):175–308
- Bonanno G, Caldarelli G, Lillo F, Mantegna RN (2003) Topology of correlation-based minimal spanning trees in real and model markets. *Phys Rev E* 68(4):046130
- Bonanno G, Caldarelli G, Lillo F, Miccichè S, Vandewalle N, Mantegna RN (2004) Networks of equities in financial markets. *Eur Phys J B* 38:363–371
- Borgatti S, Mehra A, Brass D, Labianca G (2009) Network analysis in the social sciences. *Science* 323(5916):892–895
- Boss M, Elsinger H, Summer M, Thurner S (2004) The network topology of the interbank market. *Quant Finance* 4:677–684
- Bougheas S, Kirman AP (2014) Complex financial networks and systemic risk: a review. CESifo Working Paper Series 4756, CESifo Group, Munich
- Braverman A, Minca A (2014) Networks of common asset holdings: aggregation and measures of vulnerability. SSRN: <http://ssrn.com/abstract=2379669>
- Brookfield D, Boussabaine H, Su C (2013) Identifying reference companies using the book-to-market ratio: a minimum spanning tree approach. *Eur J Finance* 19(6):466–490
- Caldarelli G (2008) *Scale free networks*. Oxford University Press, ISBN: 9780199211517
- Caraiani P (2013) Using complex networks to characterize international business cycles. *PLoS One* 8(3), e58109. doi:[10.1371/journal.pone.0058109](https://doi.org/10.1371/journal.pone.0058109)
- Cetorelli N, Peristiani S (2013) Prestigious stock exchanges: a network analysis of international financial centers. *J Bank Finance* 37(5):1543–1551
- Chapelle A, Szafarz A (2005) Controlling firms through the majority voting rule. *Phys A* 355(2):509–529
- Chapelle A, Szafarz A (2007) Control consolidation with a threshold: an algorithm *IMA J Manag Math* 18(3):235–243
- Cohen L, Frazzini A, Malloy C (2008) The small world of investing: board connections and mutual fund returns. *J Polit Econ* 116(5):951–979
- Cohen L, Malloy C, Frazzini A (2010) Sell side school ties. *J Finance* 65:1409–1437
- Cont R, Moussa A, Santos EB (2013) Network structure and systemic risk in banking systems. No. hal-00912018
- Croci E, Grassi R (2014) The economic effect of interlocking directorates in Italy: new evidence using centrality measures. *Comput Math Organ Theory* 20(1):89–112
- D'Arcangelis AM, Rotundo G (2014) Mutual funds relationships and performance analysis. *Qual Quant* 49(4):1573–1584
- D'Errico M, Grassi R, Stefani S, Torriero A (2009) Shareholding networks and centrality: an application to the Italian financial market. In: Naimzada AK, Stefani S, Torriero A (eds) *Networks, topology and dynamics: theory and applications to economics and social systems*. Lecture notes in economics and mathematical systems. Springer, Heidelberg, pp 215–228
- De Benedictis L, Tajoli L (2011) The world trade network. *World Econ* 34:1417–1454

- Del Giudice A, Marinelli N, Vitali S (2014) Sovereign wealth funds and target firms: does 'networking' matter? *J Finance Manag Mark Inst* 2:185–206
- Elton EJ, Gruber MJ, Brown SJ, Goetzmann WN (2014) *Modern portfolio theory and investment analysis*, 9th edn
- Flath D (1992) Indirect shareholding within Japan's business groups. *Econ Lett* 38:223–227
- Gabbi G, Germano G, Hatzopoulos V, Iori G, Politi M (2012) Market microstructure, bank's behaviour and interbank spreads (report no. 12/06). Department of Economics, City University London
- Gai P, Kapadia S (2010) Contagion in financial networks. *Proc R Soc A*. doi:10.1098/rspa.2009.0410
- Gai P, Haldane A, Kapadia S (2011) Complexity, concentration and contagion. *J Monet Econ* 58(5):453–470
- Garas A, Argyrakis P, Rozenblat C, Tomassini M, Havlin S (2010) Worldwide spreading of economic crisis. *New J Phys* 12:113043
- Garas A, Schweitzer F, Havlin S (2012) A k-shell decomposition method for weighted networks. *New J Phys* 14, 083030
- Garlaschelli D, Loffredo MI (2004) Fitness-dependent topological properties of the World Trade Web. *Phys Rev Lett* 93
- Gligor M, Ausloos M (2007) Cluster structure of EU-15 countries derived from the correlation matrix analysis of macroeconomic index fluctuations. *Eur Phys J B* 57:139–146
- Gligor M, Ausloos M (2008) Convergence and cluster structures in EU area according to fluctuations in macroeconomic indices. *J Econ Integr* 23:297–330
- Gower JC (1986) Metric and Euclidean properties of dissimilarity coefficients. *J Classif* 3:5–48
- Grassi R (2010) Vertex centrality as a measure of information flow in Italian Corporate Board Networks. *Phys A* 389(12):2455–2464
- Grassi R, Patarnello A, Szpilska V (2008) Corporate board network and information flows in the Italian Stock Exchange (short paper). In: *Proceedings of methods, models and information technologies for decision support systems (MTISD)*, Università del Salento, pp 110–112
- Guo W, Minca A, Wang L (2015) The topology of overlapping portfolio networks. Available at SSRN: <http://ssrn.com/abstract=2619514>
- Halaj G, Kok C (2015) Modeling emergence of the interbank networks. *Quant Finance* 15:653–671
- Leitner Y (2005) Financial networks: contagion, commitment, and private sector bailouts. *J Finance* 60(6):2925–2953
- López-Pintado D (2006) Contagion and coordination in random networks. *Int J Game Theory* 34(3):371–381
- Lucarelli C, Molyneux P, Vitali S (2012) Network features of European trading venues. Available at SSRN: <http://ssrn.com/abstract=1981843> or <http://dx.doi.org/10.2139/ssrn.1981843>
- Maeng SE, Choi HW, Lee JW (2012) Complex networks and minimal spanning trees in international trade networks. *Int J Mod Phys Conf Ser* 16:51–60
- Mantegna RN (1999) Hierarchical structure in financial markets. *Eur Phys J B* 11:193–197
- Markose S, Giansante S, Shaghghi AR (2012) Too interconnected to fail financial network of US CDS market: topological fragility and systemic risk. *J Econ Behav Organ* 83(3):627–646
- Markowitz H (1952) Portfolio selection. *J Finance* 7(1):77–91
- Miskiewicz J, Ausloos M (2006) G7 country gross domestic product (GDP) time correlations. A graph network analysis. In: Takayasu H (ed) *Practical fruits of econophysics*. Springer, Tokyo, pp 312–316
- Miskiewicz J, Ausloos M (2010) Has the world economy reached its globalization limit? *Phys A* 389:797–806
- Nature, focus issue on Complex networks in finance (2013) <http://www.nature.com/nphys/focus/finance/index.html>
- Naylor M, Rose L, Moyle B (2007) Topology of foreign exchange markets using hierarchical structure methods. *Phys A* 382:199–208
- Newman MEJ, Barabasi AL, Watts DJ (2006) *The structure and dynamics of networks*. Princeton University Press, Princeton

- Pantaleo E, Tumminello M, Lillo F, Mantegna RN (2011) When do improved covariance matrix estimators enhance portfolio optimization? An empirical comparative study of nine estimators. *Quant Finance* 11(7):1067–1080
- Pastor-Satorras R, Rub M, Diaz-Guilera A (eds) (2003) *Statistical mechanics of complex networks*. Springer, Berlin
- Pozzi F, Aste T, Rotundo G, Di Matteo T (2008) Dynamical correlations in financial systems. In: Abbott D, Aste T, Bachelor M, Dewar R, Di Matteo T, Guttman T (eds) *Complex systems II*. Proceedings SPIE 6802, 68021E. ISBN 978-1-60560-322-3
- Rossi AG, Blake D, Timmermann A, Tonks I, Wermers R (2015) Network centrality and pension fund performance. CFR working papers no 15-16, University of Cologne, Centre for Financial Research
- Rotundo G (2011) Centrality measures in shareholding networks. In: Duman E, Atiya A (eds) *Use of risk analysis in computer-aided persuasion*. NATO Science for Peace and Security Series E: Human and Societal Dynamics, 88, pp. 12–28. ISBN 978-1-60750-827-4 (print) ISBN 978-1-60750-828-1 (online), ISSN 1874-6276
- Rotundo G, D’Arcangelis AM (2010a) Ownership and control in shareholding networks. *J Econ Interact Coord* 5(2):191–219
- Rotundo G, D’Arcangelis AM (2010) Network analysis of ownership and control structure in the Italian Stock market. *Advances and applications in statistical sciences*, ISSN 0974-68119, Special Issue 2, pp 255–273
- Rotundo G, D’Arcangelis AM (2014) Network of companies: an analysis of market concentration in the Italian stock market. *Qual Quant* 48(4):1893–1910
- Sandoval L Jr, De Paula Franca I (2012) Correlation of financial markets in times of crisis. *Phys A* 391(1–2):187–208
- Schweitzer F, Fagiolo G, Sornette D, Vega-Redondo F, Vespignani A (2009) Economic networks: the new challenges. *Science* 325(5939):422
- Solis R (2009) Visualizing stock-mutual fund relationships through social network analysis. *Global J Finance Bank Issues* 3(3):8–22
- Steinbacher M, Steinbacher M, Steinbacher M (2013) Credit contagion in financial markets: a network-based approach. Available via SSRN. <http://papers.ssrn.com/sol3/papers.cfm?Abstractid=2068716>
- Varela Cabo LM, Rotundo G, Ausloos M, Carrete J (2015) Complex networks analysis in socioeconomic models. In: Commendatore P, Kayam SS, Kubin I (eds) *Complexity and geographical economics – topics and tools*. Springer series; Dynamic modeling and econometrics in economics and finance, 19. ISBN 978-3-319-12805-4, pp 209–245
- Vitali S, Glattfelder JB, Battiston S (2011) The network of global corporate control. *PLoS One*. doi:10.1371/journal.pone.0025995
- Vitting Andersen J, Nowak A, Rotundo G, Parrott L, Martinez S (2011) Price-quakes shaking the world’s stock exchanges. *PLoS One* 6(11):e26472. doi:10.1371/journal.pone.0026472
- White L (2014) The basics of “Too Big to Fail”. In: Schultz PH (ed) *Perspective in Dodd-Frank and Finance*. MIT Press