Chapter 7
The Structure of Financial Networks

Stefano Battiston, James B. Glattfelder,
Diego Garlaschelli, Fabrizio Lillo,
and Guido Caldarelli

Abstract We present here an overview of the use of networks in Finance and Economics. We show how this approach enables us to address important questions as, for example, the structure of control chains in financial systems, the systemic risk associated with them and the evolution of trade between nations. All these results are new in the field and allow for a better understanding and modelling of different economic systems.

7.1 Introduction

The use of network theory in financial systems is relatively recent, but it has exploded in the last few years, and since the financial crisis of 2008–2009 this topic

S. Battiston (✉) · J.B. Glattfelder
Systemgestaltung, ETH-Zentrum, Zurich, Switzerland
e-mail: sbattiston@ethz.ch

J.B. Glattfelder
e-mail: jglattfelder@ethz.ch

D. Garlaschelli
Said Business School, University of Oxford, Oxford, UK

F. Lillo
Dipartimento di Fisica e Tecnologie Relative, Università di Palermo, Palermo, Italy
e-mail: lillo@unipa.it

F. Lillo
Santa Fe Institute, Santa Fe, USA

G. Caldarelli
Centre SMC, and ISC CNR, Dip. Fisica, University "Sapienza", Piazzale Aldo Moro 5, Rome, Italy
e-mail: Guido.Caldarelli@roma1.infn.it

E. Estrada et al. (eds.), *Network Science*,
DOI 10.1007/978-1-84996-396-1_7,

has received even more interest. Many approaches have been attempted, and in this chapter we try to summarise the most interesting and successful of them. This is only one aspect of the increasing interest in real-world complex networks, indeed some previous studies revealed unsuspected regularities such as scaling laws which are robust across many domains, ranging from biology or computer systems to society and economics [40, 63, 68]. This has suggested that universal, or at least generic, mechanisms are at work in the formation of many such networks. Tools and concepts from statistical physics have been crucial for the achievement of these findings [13, 28]. From the initial activity, the research in the field of networks evolved so that three levels of analysis are nowadays possible. The first level corresponds to a purely topological approach (best epitomised by a binary adjacency matrix, where links simply exist or not). Allowing the links to carry weights [5], or weights and direction [66], defines a second level of complexity. Only recent studies have started focusing on the third level of detail in which the nodes themselves are assigned a degree of freedom, sometimes also called fitness. This is a non-topological state variable which shapes the topology of the network [23, 41, 43]. In this chapter, we shall try to cover these three aspects by mainly focussing on the methodological aspects, i.e. which types of networks can be constructed for financial systems, and evaluating the empirical results on networks obtained by investigating large databases of financial data ranging from individual transactions in a financial market to strategic decisions at a bank level. Readers have to note that edges and vertices can assume different meanings according to the system considered. In some situations, we can have different details of data available, and cannot investigate the three levels of detail in the same way. The cases of study we want to present here are those related to

- Networks to extract information from *Correlation Matrices*
- Networks of control as, for example, the *Ownership network* and the *Board of Directors* network
- Trading networks as the *World Trade Web* and the *Banks' Credit* networks

Since these fields can be very different, it is important to divide and classify the networks that we could encounter. An important classification of networks in general, and that will be considered here, is the one that divides networks into similarity based networks and direct interaction networks. To be specific, consider a network whose nodes are financial agents: investors, banks, hedge funds, etc. They differ in the meaning of the links. In similarity-based networks, we draw a link when the two vertices share some feature: strategy, behaviour, income, etc. In this case, one needs to assign a criterion to establish whether the similarity between two agents is relevant enough that we can join the agents with a link. This means that the agents do not interact with each other, but can be connected if they are similar. Conversely, in direct interaction networks a link between two nodes signals the presence of an interaction between the entities represented by the two nodes connected by the link. In the financial case, the interaction can be a transaction between two agents, an ownership relation of one node with respect to another, a credit relation, etc. All these situations are present in real financial networks and we shall analyse various instances of them. The structure of this chapter reflects the above division of topics so that, Sect. 7.2 overviews the studies of networks in a data mining framework (mostly similarity-based networks). In Sect. 7.3, we analyse the social aspects of the control networks where vertices are real agents or companies, and finally in Sect. 7.4, we analyse some of the most important "trading" webs.

7.2 Similarity-Based Networks

In similarity based networks, a weighted link between two nodes represents a similarity (but not necessarily a direct interaction) between the two nodes. Let us consider a system composed of N elements. Each element is represented by T variables. These variables describe several different properties of the elements (or they can represent values of the variables at different times as it is the case of time series). Typically, one introduces a matrix C describing the various $N \times N$ similarity measurements. We indicate with c_{ij} the generic element of the matrix C. In the language of networks, the matrix C identifies a complete (all edges drawn) weighted network where every of the N elements is represented by a node, and the link connecting nodes i and j is associated to a weight related to c_{ij}. From a purely topological point of view, this does not produce anything usable. In fact, if, for example, we consider the similarities between 1000 different elements, one has to check almost one million entries. The natural choice is then to exploit the extra information embedded in the values of the weights. In other words, it is necessary to filter out the most relevant links of the network. The different methods to perform this filtration give rise to different similarity based networks.

The choice of the similarity measure is also arbitrary. A very common choice is the use of the linear correlation as a measure of similarity. In this case, we have the correlation based networks. The best example of similarity-based networks are correlation based networks. In this case, the similarity between two elements (nodes) of the system is quantified by the linear correlation. The matrix element c_{ij} is the linear (or Pearson's) cross-correlation between element i and j, i.e.

$$c_{ij} = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{\langle r_i^2 \rangle - \langle r_i \rangle^2} \sqrt{\langle r_j^2 \rangle - \langle r_j \rangle^2}} \tag{7.1}$$

where r_i and r_j are the investigated variables and the symbol $\langle \ldots \rangle$ is a statistical average defined as

$$\langle r_i \rangle = \frac{1}{T} \sum_{k=1}^{T} r_i(k) \tag{7.2}$$

where $r_i(k)$ is the kth variable of element i. The correlation coefficient has values between -1 and $+1$, corresponding to perfectly anticorrelated and perfectly correlated (i.e. identical) variables, respectively. When the coefficient is zero the variables are not correlated. The correlation coefficient can be associated to a Euclidean distance with the relation [48, 59]

$$d_{ij} = \sqrt{2(1 - c_{ij})}. \tag{7.3}$$

Finally, it is important to stress that correlation coefficients are statistical estimators and therefore are subject to measurement noise. This means, for example, that two variables could be uncorrelated, but their sample correlation coefficient measured over a set of T variables is typically different from zero. Moreover, since one is interested simultaneously in many variables, another common statistical problem, named the curse of dimensionality, is common in correlation based networks. Coming back to our previous example of the correlation matrix, we have $N(N-1)/2$ distinct elements to be estimated. We remind that each element is represented by T variables so that the number of data points for this estimation is NT. Therefore, unless $T \ll N$, the statistical reliability of the similarity matrix is small because we have a small number of independent measurements per estimated parameter. Therefore, one has to devise a method to reduce the number of variables from the large N^2 value to a subset statistically more significant.

There are many different methods to construct correlation (or similarity) based networks. In this paper, we shall consider two classes, namely threshold networks and hierarchical networks, and we discuss the application of these methods to financial networks.

7.2.1 Threshold Methods

The simplest filtered network is a threshold network. Given the set of correlation or even distances between the vertices, one keeps only those above a determined level of confidence. A suitable choice, based on statistical consideration, could be the following. If the N variables are described by independent Gaussian time series of length T and we use the above defined correlation measure, it is known that for large T the distribution of the sample correlation coefficient of two uncorrelated Gaussian variable can be approximated by a Gaussian distribution with zero mean and standard deviation equal $1/\sqrt{T}$. In this case, a reasonable approach could be to consider a threshold of three standard deviations and to keep the edge values whose associated correlation coefficients are larger in absolute value than $3/\sqrt{T}$. In this way, only links associated with statistically significant correlation are preserved in the filtered network.

Another method for choosing the threshold is to randomise the data by permutation experiments and preserve in the original networks only those links associated to a correlation which is observed in the randomised data with a small probability. In this way, no a priori hypothesis on the probability distribution of data is made, and this approach is very useful especially when data distribution is very different from a Gaussian.

The networks generated with threshold methods are, in general, disconnected. If the system presents a clear cluster organisation, threshold methods are typically able to detect them. On the contrary, if the system has a hierarchical structure, threshold methods are not optimal, and hierarchical methods described below should be preferred.

7.2.2 Hierarchical Methods

A different class of filtering methods are the hierarchical methods. These are devised to detect specifically the hierarchical structure of the data. A system has a hierarchical structure when the elements of the system can be partitioned into clusters which, in turn, can be partitioned into subclusters, and so on up to a certain level. In multivariate statistics, there is a large literature on hierarchical clustering methods [2]. These methods produce dendrograms to describe the hierarchical structure of the system. One interesting aspect for network theory is that hierarchical clustering methods are also often associated with networks which are different from the dendrograms. Here we review the use of these networks as similarity based networks. For a review of the hierarchical methods and their application in finance, see also [76].

One of the most common algorithms to detect a possible hierarchical structure hidden in the data is given by the Minimum Spanning Tree (MST) procedure. The MST is the spanning tree of shortest length. In our case, the length of a link is inversely related to the similarity between the nodes connected by the link. Thus one can either use a relation as the one in (7.3) to obtain a length measure of the links and find the spanning tree of minimum length, or alternatively consider the similarity (or correlation) as "length" and look for the spanning tree of maximum length.

There are several algorithm to extract the MST. A very intuitive one is the following procedure.

- Assign distances between the vertices in such a way that the largest is the correlation between two vertices, while the shortest is the distance.
- Rank these distances from the shortest to the longest.
- Start from the shortest distance and "draw" the edge between the vertices.
- Iterate this procedure until you find an edge that would form a loop. In this case, jump to the next distance (if necessary repeat this operation).
- Stop when all the vertices have been considered.

The resulting graph is the MST of the system and the connected components progressively merging together during the construction of the graph are clusters progressively merging together. The MST is strongly related with a well known hierarchical clustering algorithm, called Single Linkage Cluster Analysis. Recently it has been shown that another network, termed Average Linkage Minimum Spanning Tree, can be associated with the most common hierarchical clustering algorithm, the Average Linkage Cluster Analysis [78].

By using different constraints, it is possible to define other hierarchical networks. The idea is that from the same correlation matrix one can also obtain correlation based networks having a structure more complex than a tree. In the case of Planar Maximally Filtered Graph (PMFG) [3, 77], we can allow loops and cliques by modifying the above algorithm. In particular, you can continue drawing edges, provided that the graph remains planar. That is to say, it should be possible to draw the graph without having to cross two different edges. Since the maximum complete subgraph one can draw with such a feature is K^4, one cannot have cliques of size larger than four in a PMFG. Authors of [3, 77] introduce a correlation based graph obtained by connecting elements with largest correlation under the topological constraint of fixed genus $G = 0$. The genus is a topologically invariant property of a surface defined as the largest number of nonintersecting simple closed curves that can be drawn on the surface without separating it. Roughly speaking, it is the number of holes in a surface. A basic difference of the PMFG with respect to the MST is the number of links which is $N - 1$ in the MST and $3(N - 2)$ in the PMFG. Moreover, the PMFG is a network with loops, whereas the MST is a tree. It is worth recalling that in [77] it has been proven that the PMFG always contains the MST. Thus the PMFG contains a richer structure than the MST (which is constrained by the tree requirement), but the number of links is still $O(N)$ rather than $O(N^2)$ as for the complete graph. Therefore, the PMFG is an intermediate stage between the MST and the complete graph. In principle, this construction can be generalised to a genus $G > 0$, i.e. one adds links only if the graph can be embedded in a surface with genus G.

7.2.3 An Application to NYSE

The main application of similarity based networks to finance has been the characterisation of the cross-correlation structure of price returns in a stocks portfolio. In this case, the nodes represent different stocks or, in general, different assets. The T variables characterising a node are the time series of price return of the associated stock computed over a given time horizon (typically one day, i.e. daily returns). The cross-correlation between stock returns measures how much the price of the two stocks move in a similar way. The characterisation and filtering of the stock return correlation matrix is of paramount importance in financial engineering and specifically in portfolio optimisation.

The threshold method selects the highest correlations and therefore considers the links in the network that are individually more relevant from a statistical point of view. An example of the analysis of the financial correlation matrix with a threshold network was given in [64, 65]. Here the authors computed the price return correlation matrix of a portfolio of $N \simeq 100$ US stocks. The variable considered is the daily price return and they constructed the correlation matrix of these N stocks. The process starts from an empty graph with no edges where the N vertices correspond to stocks. Then they introduced one edge after another between the vertices according to the rank of their correlation strength (starting from the largest value). These "asset graphs" can be studied as the number of links grows when one lowers the threshold. Typically, after a small number of edges (20 for this case study), several cycles appear in the asset graph. As more edges are added, the existing clusters are significantly reinforced and the graph remains disconnected. The graph becomes completely connected, only after a large number of edges are added (typically 1000 for a network of 116 nodes). It is possible to check that even for moderate values of the number of edges added, the connected components (clusters) are quite homogeneous in terms of the industrial sector of the stocks involved. It is also possible to study how the mean clustering coefficient changes when the links are added. Interestingly, three different regimes appear. In the first regime, we have a rapid growth of the clustering coefficient corresponding to an addition of the top 10% of the edges. In other words, the first 10% of the edges add substantial information to the system. In the second regime, when the fraction of added edges is between 10% and 20%, the rate of change of the clustering coefficient starts to slow down and reaches a sort of plateau. Finally, the edges added in the last regime, i.e. having the correlation coefficients with a rank below 30%, have relatively poor information content, and are possibly just noise.

Concerning hierarchical methods, Rosario Mantegna [58, 59] was one of the first to apply this procedure to the price return of a portfolio of stocks. Since this type of analysis has been performed afterwards in many different markets, time periods, and market conditions, the results described below are quite general. At a daily time scale, the MST is able to identify quite clearly the economic sectors of the stocks in the portfolio. Each company belongs to an economic sector (such as Basic Materials, Energy, Financial, Technology or Conglomerates) describing the core activity of the company and assigned by some external company (such as Forbes). The MST shows the presence of clusters of nodes (stocks) which are quite homogeneous in the economic sector. Often the MST displays also a structure in subclusters where nodes are stocks mostly belonging to the same subsector (for example, Communication services and Retail are subsectors of the sector of Services). In this ability of identifying the economic sectors from the correlation matrix, the MST performs typically pretty well when compared with other more traditional methods, such as spectral methods [17]. In this latter procedure, one extracts the eigenvectors of the correlation matrix and identifies sectors as groups of stocks which have a large component (compared to the others) in an eigenvector. Despite the fact that this method gives some useful information [47], the eigenvectors sometimes mix different economic sectors (especially when the eigenvalues are close to each other). A second feature typically found in the MST of daily returns of a portfolio of stocks is the presence of few nodes (stocks) with a very large degree. A typical example is observed in a portfolio of US stocks in the late 1990s. In this case, it has been found that General Electric is a hub in the network, connecting different parts of the tree associated with different economic sectors, but also connecting to leaves, i.e. stocks that are connected to the tree only trough General Electric. General Electric is a conglomerate company, and in the 1990s it was the most capitalised stock in the US financial markets. A conglomerate company is a combination of more firms engaging in entirely different businesses. These two aspects explain why General Electric was a hub of the network. The fact that hubs are often conglomerates is explained by the fact that a conglomerate company tends to be less affected by industry specific factors and more affected by global market behaviour.

It is interesting to note that when one considers price returns computed on intra-day time horizons, the topological structure of the MST changes dramatically [9]. It has been observed that on very short time horizons the MST has roughly a star like structure with the notable exception of the presence of a cluster of technological stocks. As the time horizon used to compute returns is increased, the MST becomes more structured and clusters of stocks belonging to the same economic sector progressively emerge. It is worth noting that the time horizon needed, because a given sector is evident as a cluster in the MST, depends on the considered sector. This has been interpreted as an indication that the market "learns" its cluster structure progressively.

A direct comparison of the MST and an asset graph with $N-1$ edges (like the MST) shows that the two graphs share only 25% of the edges, i.e. that 75% of the $N-1$ strongest correlation are not represented in the MST. Therefore, while the MST can inform about the taxonomy of a market, the graph asset built in this way seems to better represent the strongest correlations.

7.2.4 Other Similarity Based Hierarchical Networks in Finance

Hierarchical networks (mainly the MST) have been applied to several other financial variables. These, for example, include: (i) time series of price volatility, (ii) interest rates, (iii) hedge funds' net asset values, and (iv) trading activity of agents. In the first case [60], the investigated variable is the stock price volatility which is a key economic variable measuring the level of price fluctuations, and the MST resembles most of the properties of the price return MST. The interest rates are important economic variables, and the empirical analysis [26] shows that interest rates cluster in the MST and in the PMFG according to their maturity dates. The application of MST to hedge funds' net asset value [61] allows identifying clusters of funds adopting similar strategies (which sometimes are different from the strategies declared by the fund). Finally, the emerging study of networks of agents trading in the market [30] shows that the correlation matrix of the inventory variation of the market members trading a stock in the Spanish Stock Market contains information on the community structure of the investors.

7.3 Control Networks: The Case of Directors and Ownerships

In this case of study, the networks are used to detect chains of control that could be hidden in a traditional analysis. Here the vertices are mainly financial agents, both in the case of Board of Directors where people investigate the relation between different persons and in the case of ownership networks where people want to know about the chains of control behind a particular stock. The physics literature on complex economic networks has previously focused on both of these topics (for the boards of directors, see, for example, [6, 62], while for market investments, see [8, 41]). At the same time, also in economics there is a vast body of literature on corporate control that focuses on corporations as individual units. The research topics this field of study addresses can be grouped into three major categories: firstly, analysing the dispersion or concentration of control [29, 73]; secondly, empirically investigating how the patterns of control vary across countries and what determines them [55, 56]; and thirdly, studying the impact of frequently observed complex ownership patterns [11, 15, 16, 34] such as the so-called pyramids [1] and cross-shareholdings (also known as business groups) [49].

In addition, research in cooperative game theory analysing political voting games has resulted in the development of the so-called power indices [4, 72]. These ideas have been applied to coalitions of shareholders voting at Shareholders Meetings [57].

7.3.1 Stock Ownership Network

It should be noted that most previous empirical studies did not build on the idea that ownership and control define a vast complex network of dependencies. Instead, they selected samples of specific companies and looked only at their local web of interconnections. These approaches are unable to discern control at a global level. This emphasises the fact that the bird's-eye-view given by a network perspective is important for unveiling overarching relationships. Remarkably, the investigation of the financial architecture of corporations in national or global economies taken *as a whole* is just at the beginning [18, 41, 54]. Here we present an analysis based on [45] that allows us to show how control is distributed at the country level, based on the knowledge of the ownership ties.

The dataset considered spans 48 countries and is compiled from Bureau van Dijk's ORBIS database.[1] In this subset, there are a total of 24,877 stocks and 106,141 shareholding entities who cannot be owned themselves (individuals, families, cooperative societies, registered associations, foundations, public authorities, etc.). Note that because the corporations can also appear as shareholders, the network does not display a bipartite structure. The stocks are connected through 545,896 ownership ties to their shareholders. The database represents a snapshot of the ownership relations at the beginning of 2007. The values for the market capitalisation, which is defined as the number of outstanding shares times the firm's market price, are also from early 2007. These values will be our proxy for the size of corporations and hence serve as the non-topological state variables.

The network of ownership relations in a country is very intricate and a cross-country analysis of some basic properties of these networks reveals a great level of

[1] http://www.bvdep.com/orbis.html.

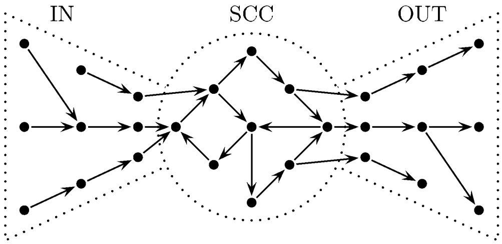


Fig. 7.1 Schematic illustration of a bow-tie topology: the central area is the strongly connected component (SCC) where there is a path from each node to every other node and the *left* (IN) and *right* (OUT) sections contain the incoming and outgoing nodes, respectively

variability. For instance, at the topological level, the analysis of the number and sizes of connected components. Many small components correspond to a fragmented capital market, while a giant and dense component corresponds to an integrated market. It is, however, not very clear what such connected components reveal about the structure and distribution of control. The same pattern of connected components can feature many different configurations of control. Therefore, it makes sense to move on to the next level of analysis by introducing the notion of direction. Now it is possible to identify strongly connected components. In terms of ownership networks, these patterns correspond to sets of corporations where every firm is connected to every other firm via a path of indirect ownership. In particular, these components may form bow-tie structures, akin to the topology of the World Wide Web [12]. Figure 7.1 illustrates an idealised bow-tie topology. This structure is particularly useful to illustrate the flow of control, as every shareholder in the IN section exerts control and all corporations in the OUT section are controlled.

In the above dataset, roughly two thirds of the countries' ownership networks contain bow-tie structures (see also [79]). Indeed, already at this level of analysis, previously observed patterns can be rediscovered. As an example, the countries with the highest occurrence of (small) bow-tie structures are KR and TW, and to a lesser degree JP (the countries are identified by their two letter ISO 3166-1 alpha-2 codes). A possible determinant is the well known existence of the so-called business groups in these countries (e.g. the *keiretsu* in JP, and the *chaebol* in KR) forming a tightly-knit web of cross-shareholdings (see the introduction and references in [49] and [33]). For AU, CA, GB and US, one can observe very few bow-tie structures of which the largest ones, however, contain hundreds to thousands of corporations. It is an open question if the emergence of these mega-structures in the Anglo-Saxon countries is due to their unique "type" of capitalism (the so-called Atlantic or stock market capitalism, see the introduction and references in [27]), and whether this finding contradicts the assumption that these markets are characterised by the absence of business groups [49].

Next to bow-tie structures, we can identify additional structures revealing the chains of control. The simplest way is by considering the core structures of the ownership networks and check the more important vertices in this subset.

The question of the vertices' importance in a network is central in graph theory, and different possibilities have been considered. Traditionally, the most intuitive quantity is the degree k, that is, the number of edges per vertex. If the edges are oriented, one has to distinguish between the in-degree and out-degree, k^{in} and k^{out}, respectively. When the edges are weighted, the corresponding quantity is called *strength* [5]:

$$k_i^{\text{w}} := \sum_j W_{ij}. \tag{7.4}$$

Note that for weighted and oriented networks, one has to distinguish between the in- and out-strengths, $k^{\text{in-w}}$ and $k^{\text{out-w}}$, respectively.

However, the interpretation of $k^{\text{in/out-w}}$ is not always straightforward for real-world networks. In the case of ownership networks, there is no useful meaning associated with these values. In order to provide a more refined and appropriate description of weighted ownership networks, we introduce two quantities that extend the notions of degree and strength in a sensible way.

The first quantity to be considered reflects the relative importance of the neighbours of a vertex. More specifically, given a vertex j and its incoming edges, we focus on the originating vertices of such edges. The idea is to define a quantity that captures the relative importance of incoming edges. When no weights are associated to the edges, we expect all edges to count the same. If weights have a large variance, some edges will be more important than others. A way of measuring the number of prominent incoming edges is to define the *concentration index* as follows:

$$s_j := \frac{(\sum_{i=1}^{k_j^{\text{in}}} W_{ij})^2}{\sum_{i=1}^{k_j^{\text{in}}} W_{ij}^2}. \tag{7.5}$$

Note that this quantity is akin to the inverse of the Herfindahl index extensively used in economics as a standard indicator of market concentration [50]. Indeed, already in the 1980s the Herfindahl index was also introduced to measure ownership concentration [19]. Notably, a similar measure has also been used in statistical physics as an order parameter [25]. In the context of ownership networks, s_j is interpreted as the effective number of shareholders of the stock j. Thus it can be interpreted as a measure of control from the point of view of a stock.

The second quantity to be introduced measures the number of important outgoing edges of the vertices. For a given vertex i with a destination vertex j, we first define a measure which reflects the importance of i with respect to all vertices connecting to j:

$$H_{ij} := \frac{W_{ij}^2}{\sum_{l=1}^{k_j^{\text{in}}} W_{lj}^2}. \tag{7.6}$$

This quantity has values in the interval (0, 1]. For instance, if $H_{ij} \approx 1$ then i is by far the most important destination vertex for the vertex j. For our ownership network,

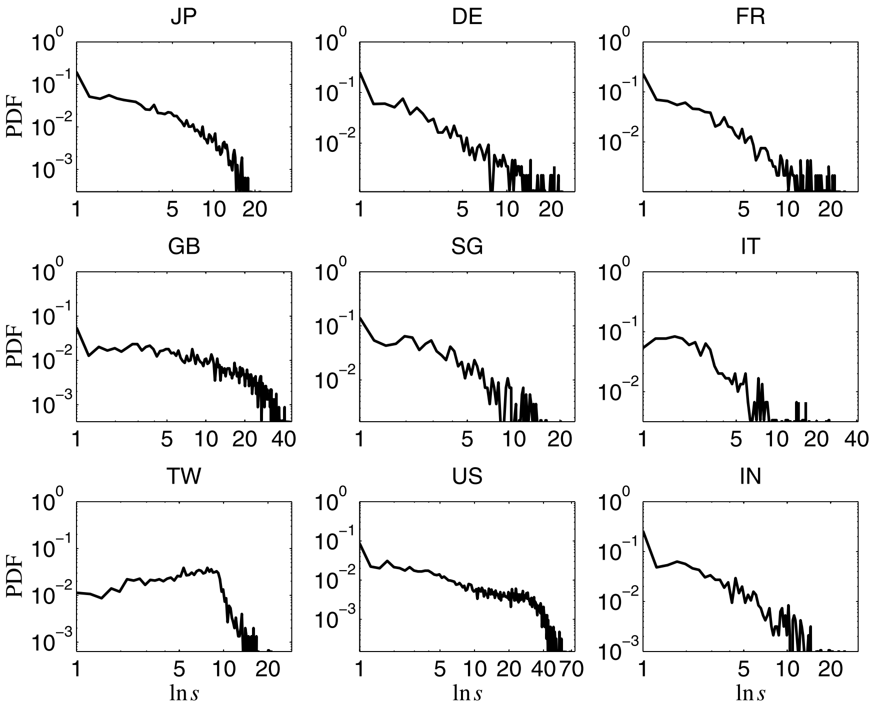


Fig. 7.2 Probability distributions of s_j for selected countries; PDF in log–log scale

H_{ij} represents the *fraction of control* shareholder i has on the company j. In a next step, we then define the *control index*:

$$h_i := \sum_{j=1}^{k_i^{\text{out}}} H_{ij}. \tag{7.7}$$

This quantity is a way of measuring how important the outgoing edges of a node i are with respect to its neighbours' neighbours. Within the ownership network setting, h_i is interpreted as the effective number of stocks controlled by shareholder i.

7.3.1.1 Distributions of *s* and *h*

The above defined measures can provide insights into the patterns of how ownership and control are distributed at a local level. In particular, Fig. 7.2 shows the probability density function (PDF) of s_j for a selection of nine countries. There is a diversity in the shapes and ranges of the distributions to be seen. For instance, the distribution of GB reveals that many companies have more than 20 leading shareholders, whereas in IT few companies are held by more than five significant shareholders. Such country-specific signatures were expected to appear due to the differences in

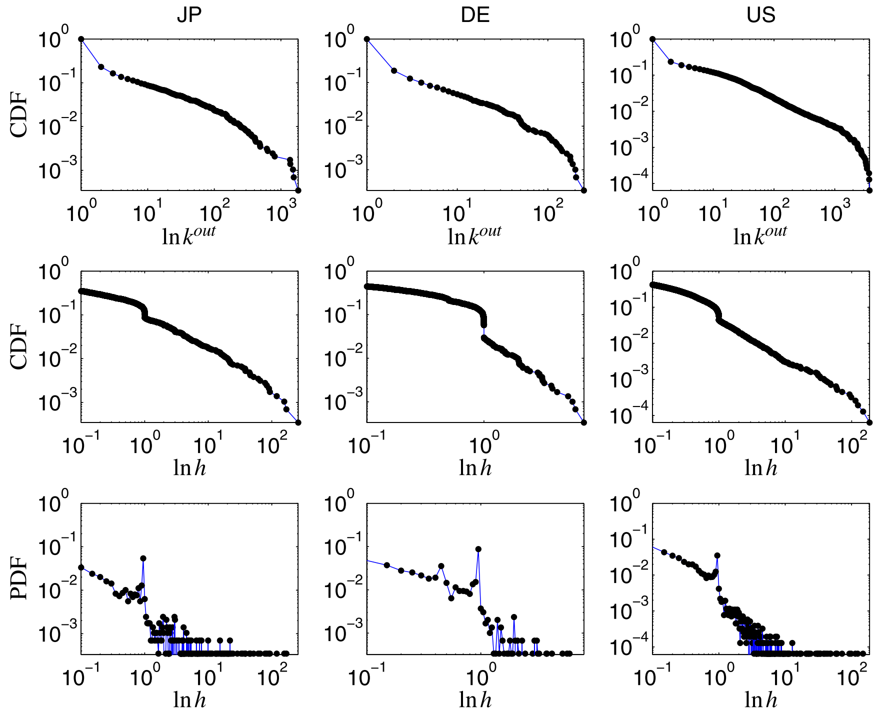


Fig. 7.3 Various probability distributions for selected countries: (*top panel*) CDF plot of k_i^{out}; (*middle panel*) CDF plot of h_i; (*bottom panel*) PDF plot of h_i; all plots are in log–log scale

legal and institutional settings (e.g. law enforcement, protection of minority shareholders [55], etc.).

On the other hand, looking at the cumulative distribution function (CDF) of k_i^{out} (shown for three selected countries in the top panel of Fig. 7.3) a more uniform shape is revealed. The distributions range between two to three orders of magnitude. Indeed, some shareholders can hold up to a couple of thousand stocks, whereas the majority have ownership in less than ten. Considering the CDF of h_i, seen in the middle panel of Fig. 7.3, one can observe that the curves of h_i display two regimes. This is true for nearly all analysed countries, with a slight country-dependent variability. Notable exceptions are FI, IS, LU, PT, TN, TW, VG. In order to understand this behaviour, it is useful to look at the PDF of h_i, shown in the bottom panel of Fig. 7.3. This uncovers a new systematic feature: the peak at the value of $h_i = 1$ indicates that there are many shareholders in the markets whose only intention is to control one single stock.

The quantities defined in (7.5) and (7.7) rely on the direction and weight of the links. However, they do not consider non-topological state variables assigned to the nodes themselves. In our case of ownership networks, a natural choice is to use the market capitalisation value of firms in thousand USD, v_j, as a proxy for their sizes. Hence v_j will be utilised as the state variable in the subsequent analysis. In a first step, we address the question of how much wealth the shareholders own, i.e. the value in their portfolios. As the percentage of ownership given by W_{ij} is a measure of the fraction of outstanding shares i holds in j, and the market capitalisation of j is defined by the number of outstanding shares times the market price, the following quantity reflects i's *portfolio value*:

$$p_i := \sum_{j=1}^{k_i^{\text{out}}} W_{ij} v_j. \tag{7.8}$$

Extending this measure to incorporate the notions of control, we replace W_{ij} in the previous equation with the fraction of control H_{ij}, defined in (7.6), yielding the *control value*:

$$c_i := \sum_{j=1}^{k_i^{\text{out}}} H_{ij} v_j. \tag{7.9}$$

A high c_i value is indicative of the possibility to control a portfolio with a big market capitalisation value.

Ownership networks are prime examples of real-world networks, where all three levels of complexity play a role. By incorporating all this information in their empirical analysis, it is possible to discover a rich structure in their organisation and unveil novel features. Moreover, the new network measures introduced above can be applied to any directed and weighted network, where the nodes can be assigned an intrinsic value.

7.3.2 Board of Directors

Another class of financial networks tracks the structure of the control of the companies by considering the people sitting on the board of directors. This is an example of a bipartite graph, that is to say, a graph that can be divided in two nontrivial subsets such that the edges connect one element of the first set with an element in the second one. The first set of nodes is composed of the companies, while the other set is composed of the people who sit in at least one board of directors. A link between a person and a company exists if that person sits on the board of directors of that company. As in all bipartite networks, it is possible to extract two graphs. In the first case, the nodes are the companies and a link exists if two companies share at least one common director (boards network). In the second case, the nodes are the directors, and a link exists if two directors sit together on at least one board (directors network). The empirical analysis of real data (see [20] for US companies and [7] for US and Italian companies) revealed that the networks display small world properties, are assortative and clustered. Table 7.1 shows some summary statistics from the empirical analysis of these networks. These empirical results strongly indicate the presence of lobbies of directors sitting on the same boards. Moreover, the

Table 7.1 The first three columns report data for the Boards (B) network in Italy in 1986 (B 86), in Italy in 2002 (B 02), and in USA in 1999 (B US). The last three columns report the data for the Directors (D) network in the same cases (1986, 2002 in Italy and USA in 1999)

	B 86	B 02	B US	D 86	D 02	D US
N	221	240	916	2,378	1,906	7,680
M	1,295	636	3,321	23,603	12,815	55,437
Nc/N	0.97	0.82	0.87	0.92	0.84	0.89
k/kc (%)	5.29	2.22	1.57	0.84	0.71	0.79
b/N	0.736	0.875	1.08	1.116	1.206	1.304
$\langle C \rangle$	0.356	0.318	0.376	0.899	0.915	0.884
$\langle d \rangle$	3.6	4.4	4.6	2.7	3.6	3.7
r	0.12	0.32	0.27	0.13	0.25	0.27

assortative character of the network indicates that directors that sit on many boards (and are therefore very influential) tend to sit on the same boards (more than what is expected from a random null model).

7.4 Transaction Networks: Interbank Networks and Bank–Firm Networks

The most immediate application of networks to finance and economics is given whenever we have a transaction between two agents. The latter are the vertices and the transaction is the edge between them. In particular, many theoretical and empirical studies have considered the networks where the nodes are the banks and a link represents a possible (directed and weighted) relation between banks. For this system, two types of networks have been considered, the interbank market and the payment system. Boss et al. [10] were among the first to study the interbank market with concepts and tools of modern network theory. They empirically studied the complete Austrian interbank market composed of around 900 banks in the years 2000–2003. The interbank market is described as a network where the vertices given by the banks are nodes and the claims and liabilities between them are described as links. The interbank market is therefore a weighted and directed network. Following a rule similar to that of food webs, the direction of the link goes from the bank having a liability to the bank claiming the liability. The weight is the amount (in money) of the liability. Authors investigated the community structure of this network by applying the Girvan–Newman algorithm based on link betweenness [44]. They found that the sectoral organisation of the banking industry (in terms of saving banks, agricultural banks, joint stock banks, etc.) is rediscovered by the community detection algorithm with a 97% precision, showing that the network structure brings information on the real sectoral organisation of the banking sector. The study of the topology of the network shows that it can be modelled as a scale-free network. In fact, both the in-degree and the out-degree probability density functions display a clear power-law behaviour with a tail exponent equal to 1.7 and 3.1, respectively. The clustering coefficient of the undirected graph is equal to 0.12 and this low value has been associated to the community organisation of the network. On the other hand, the average path length is 2.59, showing a clear small-world effect. The analysis of the interbank market is important in the framework of the quantification of systemic risk [21], the risk of collapse of an entire financial system. In fact, scale-free networks are typically robust with respect to the random breakdown of nodes and fragile with respect to intentional attack against the hubs. A proper characterisation of the interbank network topology is potentially useful in determining the criticality of the financial system. Sorämaki et al. [74] have investigated the daily network of the Fedwire R Funds Service, which is a gross settlement system operated by the Federal Reserve System (USA). Today almost 10,000 participants initiate funds transfer through Fedwire for the settlement of different operations, such as the trading of federal funds, securities transactions, real estate transactions, etc. The daily value exchanged is close to $ 3 trillion. The network describing the fund transfers in a day is clearly a directed and weighted (either in terms of value or in terms of number of transfers) network. There is typically a giant undirected connected component which accounts for almost all the nodes. Inside this component, authors studied the giant strongly connected component, which comprises all the nodes that can reach each other through a directed path. This component has typically 5,000 nodes (in the investigated year 2004) and contains roughly 80% of all the nodes in the giant undirected component. The connectivity, i.e. the probability that two nodes share a link, is very low and close to 0.3%, indicating that the network is extremely sparse. The reciprocity, i.e. the fraction of links for which there is a link in the opposite direction, is 22%, even if large links (either in terms of value or in term of volume) are typically reciprocal, probably as a result of complementary business activity or the risk management of bilateral exposure. Authors also studied the dynamics of the network and found that during days of high activity in the system, the number of nodes, links, and connectivity increases, while the reciprocity is not affected. The in-degree and out-degree of the network is power law distributed with an exponent close to two, indicating a scale-free behaviour. The network is disassortative. As in the Austrian interbank market, the average path length is between two and three, suggesting a small-world behaviour. However, differently from the previous case, the mean clustering coefficient is quite large (0.53), even if the distribution of clustering coefficients across nodes is very broad, and for one fifth of the nodes the clustering coefficient is equal to zero. There is a nice scaling relation between the out-degree of a node and its strength (defined as the total volume or value exchanged transferred by the node) with an exponent 1.2–1.9. This means that more connected nodes transact a higher volume or value than would be suggested by the degree. Finally, they consider as an important case study the change in the payment network topology due to the terrorist attack of September 11th, 2001, when some infrastructure of the system was destroyed with a clear occurrence of a liquidity crisis. They found that the network became smaller, less connected, and with a larger average path length than in a typical trading day. Another type of interbank market recently investigated in the framework of network theory is the Italian interbank money market [23, 52, 53]. In order to buffer liquidity shocks, the European Central Bank requires that on average 2% of all deposits and debts owned by banks are stored at national central banks. Given this constraint, banks can exchange excess reserves on the interbank market with the objective to satisfy the reserve requirement and in order to minimise the reserve implicit costs. Authors investigated the Italian electronic broker market for interbank deposits e-MID in the period 1999–2002. They constructed a network where the nodes are the banks. For every pair of banks i and j, there is an oriented edge from i to j, if bank j borrows liquidity from bank i. The mean number of banks (nodes) active in a day is 140 and the mean number of links is 200. As in the previous bank networks, authors found that the in- and out-degree are described by a distribution with a power law tail with an exponent between two and three. They also found that the network is disassortative and that the clustering coefficient of a node decreases with its degree as a power law. The various banks operating in a market can be divided into different groups roughly related to their size, i.e. the volume of their transaction. Smaller banks make few operations and on average lend or borrow money from larger banks. The latter ones have many connections with each other, making many transactions, while the small banks do not. A visualisation of the network in a typical day shows that the core of the structure is composed of the big banks. There is also a tendency for small banks to be mostly lenders, while large banks are more likely to be borrowers. All the above quantities can be reproduced by means of a suitable model of network growth [14]. The idea is that the vertices representing the banks are defined by means of an intrinsic character that could, for example, be their size. Edges are then drawn with a probability proportional to the sizes of the vertices involved. The community analysis of the network based on the cross-correlation of the signed traded volume confirms the presence of two main communities, one mainly composed by large banks and the other composed mainly by small banks [23]. By looking at the network dynamics, the authors of this set of studies found a clear pattern where the network degree increases and the strength decreases close to the critical days when the reserves are computed. All the above studies found that the banking system is highly heterogeneous with large banks borrowing from a high number of small creditors. Iori et al. [51] showed in an artificial market model that when banks are heterogeneous high connectivity increases the risk of contagion and systemic failure.

7.4.1 Credit Networks

A different and interesting type of financial networks describes the complex credit relation between firms and banks. The reasons for a credit relation between a bank and a firm are the following. The bank supplies credit to the firm in anticipation of interest margin, while the firm borrows money form the bank in order to financing the growth of its business. A credit relation, thus, creates a strong dependence between a firm and a bank. The credit network between firms and banks is important for the understanding and quantification of systemic risk of the economic and financial system [21]. In fact, the failure of a big firm may have a strong effect on the balance sheet of a bank and the insolvency of a firm may lead many banks to fail or to change their credit policy, by increasing the interest rates or by reducing the supply of loans. This, in turn, may lead other firms to failure, which creates more financial distress among banks and so on, in a sort of domino effect affecting the whole system. An example of such scenario happened in Japan during the 1990s. A way of studying the mutual interdependency between banks and firms is through the investigation of the credit network [22, 24, 35]. In an empirical study of the Japan credit network [24, 35], roughly 200 banks and 2, 700 publicly quoted firms were considered. The credit network is a typical case of bipartite network where the two types of nodes correspond to banks and firms and a weighted link exists between bank i and firm j if, at the considered time, there is a credit relation between i and j. The weight of the link is the amount of money firm j borrows from bank i. The average degree of firms is 8, while the average degree of banks is 120. Both the degree distributions of firms and banks are fat tailed, possibly consistent with a scale-free network, and the fitted tail exponent for the cumulative distribution is around 1 for banks and 2.5 for firms. The weights in the network are highly heterogeneous, and there is a scaling relation between the degree of a node and its strength (i.e. the total amount of money borrowed or lent by the node), indicating that the average loan is larger for a larger degree (typically, larger banks). A widely discussed topic in the economics of credit is the existence of two models of credit. In Anglo-Saxon countries, firms typical establish credit relationship with few banks, while in countries such as Japan, Germany, and Italy a firm borrows moneys from more banks. The theory of optimal number of bank relationships is quite complicated because different choices have different advantages and disadvantages. The question can be investigated empirically by measuring the participation ratio of node i

$$Y_i = \sum_j \frac{w_{ij}}{s_i} \tag{7.10}$$

where w_{ij} is the weight of the link between i and j and s_i is the strength of node i. In the case of full homogeneity, the participation ratio is the inverse of the degree of node i. In a multiple bank scenario, the participation ratio is much higher, since one bank dominates the credit of firm i. The analysis of the Japan dataset confirms that large firms tend to have a participation ratio different from the inverse of the degree, i.e. large firms tend to establish credit relations with many banks. From any bipartite network, it is possible to extract two projected networks. In the bank network, two banks are linked if they share at least one borrower, while in the firm network, two firms are linked if they have at least one common lender. The weight of the links is the number of common borrowers/lenders. The authors of [24] extracted the MST from these projected networks. They found that the bank network is characterised by large hubs, and the cluster structure of this network is largely explained by the geographic region where the banks operate. In other words, banks in the same geographic region tend to lend money to the same firms, probably from the same region. The analysis of the MST of the firms in the same economic sector shows that each firm is connected to many others and that there are no communities. This is probably due to the fact that lending is not sectoral. Moreover, in almost all the investigated sectoral networks, a large hub is observed. Finally, in [35], a linear analysis of the same dataset was performed in order to measure a set of scores related to the financial fragility. It was found that these scores were statistically significant with respect to a random network and that there were periods when the structure was stable or unstable. Drastic changes were observed in the late 1980s when the bubble started in Japan. Moreover, the set of regional banks had large values of the fragility score.

7.5 The World Trade Web

The world economy is a tightly interconnected system of mutually dependent countries, and one of the dominant channels mediating international interactions is trade. When countries are represented as vertices, and their trade relations as connections, the world trade system appears as an intricate network, known in the literature as the International Trade Network or World Trade Web (WTW). The recent advances in network theory have renewed the analysis of the trade system from a perspective that fully takes into account its global topology [31, 32, 37, 42, 43, 69–71]. Indeed, traditional macroeconomic approaches have extensively addressed the empirical patterns of trade by considering the pairwise interactions involved locally between countries, but have placed much less emphasis on the analysis of higher-order properties obtained considering indirect interactions embedded in the whole international network. On the other hand, the globalisation process of the economy highlights the importance of understanding the large-scale organisation of the WTW and its evolution.

In what follows, we present some of the recent research results obtained in the analysis of the global trade network. The WTW has many possible network representations. In particular, it can be either directed or undirected, and either weighted (by taking into account the magnitude of trade flows) or unweighted. The properties of the WTW as a weighted network (where trade volumes are explicitly taken into account) are of great importance, as link intensities are extremely heterogeneous and are found to change significantly the picture obtained in the binary case [31, 32, 69, 71]. On the other hand, if regarded as a binary network (thus ignoring the magnitude of trade links), the WTW is found to be one of the few real networks whose topology can be modelled in detail using simple ideas that have been developed recently [37, 42, 43]. This implies that, besides its importance for understanding the global economic system, the WTW is particularly interesting for network theory. By contrast, unfortunately no complete model of the weighted WTW has been proposed, at least not reaching the same level of detail as its binary counterpart. At present, the only available weighted models of trade are the so-called *gravity models* which aim at predicting the observed magnitude of trade fluxes in terms of a few explanatory factors, but fail to explain their topology. So it appears that our current understanding of the trade system requires a combination of both approaches, and that the definition of a complete and satisfactory model of the WTW is still an open problem. For this reason, since here we are particularly interested in the data-driven problem of modelling the international trade network, we briefly present gravity models first, and then report in more detail various empirical properties and models of the WTW only at a purely topological level. For weighted analyses, the reader is referred to the relevant literature [31, 32, 69, 71].

7.5.1 Gravity Models

The so-called 'gravity models', first introduced by Tinbergen [75] and then rephrased in various extended or alternative forms, are the earliest econometric models of trade. Their name originates from a loose analogy with Newton's law of gravitation, where the magnitude of the force attracting two objects carrying mass is completely determined by the two masses involved and the distance separating them. Similarly, gravity models assume that the volume of trade between any two countries can be traced back to a suitable measure of the 'mass' of the two countries (representing their intrinsic economic size) and to some 'distance' between them (representing a combination of factors determining the overall resistance to trade). The economic 'mass' of a country is customarily identified with its total Gross Domestic Product (GDP in the following). Also, geographic distance is the main factor expected to determine trade resistance. However, additional factors may be incorporated in the models, having either a positive or negative effect on the volume of trade. In the simple case where one additional factor coexists with geographical distance as an explanator of trade, the gravity model predicts that the volume of trade from country i to country j in a given year is

$$w_{ij} = \beta_0 w_i^{\beta_1} w_j^{\beta_2} d_{ij}^{\beta_3} f_{ij}^{\beta_4} \tag{7.11}$$

where w_i is the total GDP of country i in the year considered, d_{ij} is the geographic distance between countries i and j, and f_{ij} is an additional factor either favouring or suppressing trade between i and j. The various factors expected to determine w_{ij} are controlled by the associated parameters $\{\beta_0, \dots, \beta_4\}$, which are global and do not depend on the two countries involved. Usually, a Gaussian error term is added to the above equation, allowing to employ standard statistical techniques to fit the model to the empirical (nonzero) trade flows and obtain the corresponding values of $\{\beta_0, \dots, \beta_4\}$. These analyses have been applied extensively on different datasets, and the general result is that β_1 and β_2 are positive and of the same order (both close to one), confirming the expectation that larger GDPs imply more trade. Similarly, empirical analyses provide evidence for a negative value of β_3, confirming a negative effect of geographic distance on trade. Depending on the nature of the additional factor f_{ij}, the parameter β_4 can be either positive or negative. Positive effects are found to be associated to two countries being members of the same economic group, having signed a trade agreement, or sharing geographic borders. By contrast, negative effects are found in presence of embargo, trade restrictions or other conditions representing a trade friction.

Gravity models have been tested extensively on many case studies reporting trade volumes, either at a global or regional level. In general, they make very good predictions of the magnitude of nonzero trade flows. However, they do not predict zero

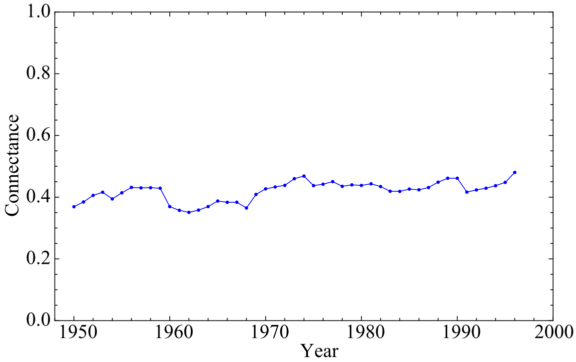


Fig. 7.4 Time evolution of the connectance c in the undirected WTW

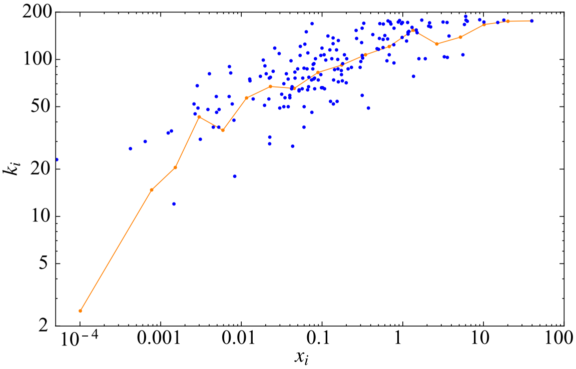
volumes, which means that they always generate a complete weighted graph where all pairs of countries exchange goods in both directions. As we now show, this is in sharp contrast with what is empirically observed. Indeed, recent analyses of the WTW have revealed that its topology is highly structured, and have thus highlighted a previously unrecognised drawback of gravity models.

7.5.2 The Heterogeneous Topology of Trade Neighbourhoods

In what follows, we show some empirical properties of the WTW topology and discuss a modelling framework that has been proposed to reproduce them. Data on the WTW topology can be obtained from the dataset [46], where annual snapshots of the network are available for all years from 1950 to 2000. We first report the properties of the WTW as an undirected network, and address its directed structure later on.

A key observation regards the heterogeneity of the number of trade partners (the degree k_i) of world countries. If only the average degree $\overline{k}$ of the WTW is considered, or equivalently the connectance $c \equiv \overline{k}/(N-1)$, then the WTW is found to display an almost constant temporal behaviour [37] (see Fig. 7.4). However, this regularity hides an intrinsic variability in the degrees. Indeed, it is found that the degree k_i of a country is positively correlated with both the total GDP and the per capita GDP of that country [43, 70]. If the total GDP w_i is considered, the relation is particularly strong [43], as shown in Fig. 7.5 for both raw and logarithmically binned data. For brevity, here and in what follows we only show the results obtained for a particular snapshot (the year 1995). These results have been shown to be robust in time through the entire period covered by the database [37]. The above result means that, on average, higher-income countries have a larger number of trade partners. And, since the total GDP is broadly distributed across world countries, this also means that one expects the number of trade partners to be broadly distributed as well. Similarly, one expects many other topological quantities, especially those involving the degree explicitly, to be strongly dependent on the GDP. As we show later on, empirical analyses have shown that this is exactly the case. These considerations

Fig. 7.5 Degree k_i as a function of the rescaled GDP $x_i \equiv w_i/\overline{w}$ in the undirected version of the WTW (year 1995): (*isolated points*) raw data; (*joined points*) logarithmically binned data



immediately imply that gravity models, which trivially produce a complete network of trade irrespective of the GDP values, are bad models of the WTW topology. This motivates the introduction of models that explicitly address the heterogeneity of trade neighbourhoods, i.e. who trades with whom in the world. We first describe one such class of models that has been proposed, before comparing its predictions to real trade data.

7.5.3 The Fitness Network Model

In the popular random graph model by Erdős and Rényi, all pairs of vertices are independently connected with probability p with an undirected edge. This is known to generate an ensemble of graphs where all vertices are statistically equivalent. For instance, the degree distribution is sharply peaked about the expected value $\langle k \rangle = p(N-1)$, where N is the number of vertices. More recently, a generalised class of random graphs, known as the *hidden variable* or *fitness* model, has been defined [14]. This model can incorporate a high degree of heterogeneity in the statistical properties of vertices. To this end, an arbitrary distribution $\rho(x)$ is specified, from which a value x_i is assigned to each vertex i. This value, interpreted as a fitness affecting the connectivity patterns of vertices differentially, is treated as a hidden variable responsible for the topological properties of the network. In particular, a link between any two vertices i and j is drawn with probability $p_{ij} = p(x_i, x_j)$. Thus $p(x_i, x_j)$ and $\rho(x)$ determine the topology of the network in a way that is simple enough to be controlled analytically, and yet allows very complex structural properties to be generated [14].

When applied to the WTW, the fitness model is a minimal model that allows to retain the basic ingredient of gravity models, i.e. that the GDP is the main factor determining trade, to the different purpose of modelling the topology, rather than the magnitude, of trade flows. Indeed, as we discussed, the degree of a vertex in the WTW is empirically found to be dependent on the GDP of the corresponding country. This finding can be exploited by defining a fitness model where the hidden variable x_i is taken to be a function of the total GDP of country i, that we denoted as w_i [43]. This means $x_i = f(w_i)$, and implies that the distribution $\rho(x)$ is no longer arbitrary, but empirically accessible. Still, one needs to specify the connection probability $p(x_i, x_j)$. The simplest nontrivial choice is one where x_i controls the degree k_i, but not other properties directly. This does not mean that other topological properties (besides k_i) will not depend on x_i, but rather that they will depend on x_i as a result of their being dependent on k_i, which is constrained directly by x_i. In more formal words, this scenario in one where the network generated by $p(x_i, x_j)$ has a specified degree sequence $\{k_i\}$ (determined by $\{x_i\}$), and is maximally random otherwise. It is known [67] that this scenario corresponds to the particular choice

$$p(x_i, x_j) = \frac{\delta x_i x_j}{1 + \delta x_i x_j} \tag{7.12}$$

where δ is a global parameter controlling the expected total number of links, and the $\{x_i\}$ distribute these links among vertices heterogeneously. The above considerations result in a model for each yearly snapshot of the WTW, where two countries are connected with probability given by (7.12), where $x_i = f(w_i)$ [43]. By assuming the simplest form of dependence, i.e. a linear relation, and reabsorbing the coefficient of proportionality in δ, then for a given year one can define x_i as an adimensional rescaled variable representing the GDP relative to the average GDP in the same year:

$$x_i \equiv \frac{w_i}{\sum_{j=1}^{N} w_j / N} \tag{7.13}$$

(note that the number of independent world countries N depends on the particular year considered). This specifies the model completely, and leaves δ as the only free parameter to be tuned in each snapshot of the network. By applying the maximum likelihood principle to the model [39], one can show that the optimal choice is the value δ^* such that

$$L = \sum_i \sum_{j<i} \frac{\delta^* x_i x_j}{1 + \delta^* x_i x_j} \tag{7.14}$$

where L is the observed number of links in that particular snapshot of the network.

After this simple parameter choice is made, one can check the predictions of the model against real data [43]. In particular, one can obtain the explicit dependence of the degrees on the GDP predicted by the model. The average degree of a vertex having fitness x_i is simply

$$\langle k_i \rangle = \sum_{j \neq i} p_{ij} \tag{7.15}$$

which is an increasing function of x_i, since p_{ij} is an increasing function of both x_i and x_j. The resulting predicted curve [43] is shown in Fig. 7.6, where the empirical (logarithmically binned) trend shown previously in Fig. 7.5 is reported again.

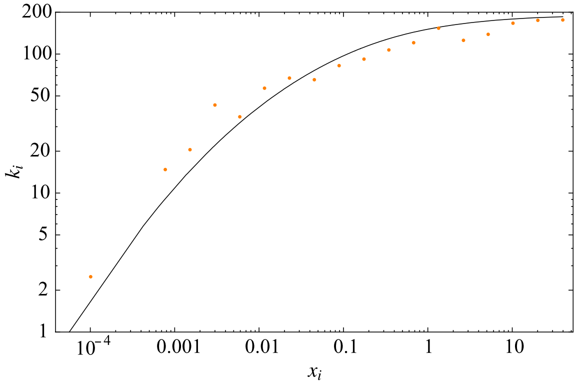


Fig. 7.6 Degree k_i as a function of the rescaled GDP $x_i \equiv w_i/\overline{w}$ in the undirected version of the WTW (year 1995): (*points*) real data (logarithmically binned); (*solid line*) theoretical prediction of the fitness model

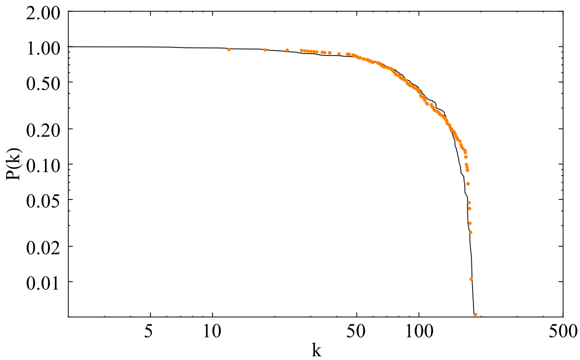


Fig. 7.7 Degree distribution in the undirected WTW (year 1995): (*points*) real data; (*solid line*) theoretical prediction of the fitness model

The agreement between empirical data and the model is very good. One can thus proceed in testing the prediction of the model against other empirical topological properties. As mentioned, the degree distribution of the WTW is found to be highly heterogeneous [43, 70]. We now describe it in more detail and show it in Fig. 7.7. One can see that there is indeed a huge variability in the number of trade partners of world countries. However, unlike other real-world networks, this variability is not captured by a scale-free distribution, as one observes an accumulation of degrees close to the maximum possible value $N-1$, which results in a strong cut-off in the right tail of the distribution [43]. The reason for this cut-off is the saturation effect shown in Fig. 7.6, i.e. the convergence of k_i to values close to $N-1$ when x_i increases. This behaviour is well reproduced by the model, and indeed the predicted degree distribution (which is shown in Fig. 7.7 superimposed to the empirical one) is in very good accordance with the data.

Additional properties of the WTW include higher-order patterns. An important one is the correlation between the average degree k_i^{nn} of the neighbours of vertex i and the degree k_i. This property is shown in Fig. 7.8. The observed decreasing trend means that, on average, highly connected countries trade with poorly connected countries, and that trade between countries of the same level of connectivity

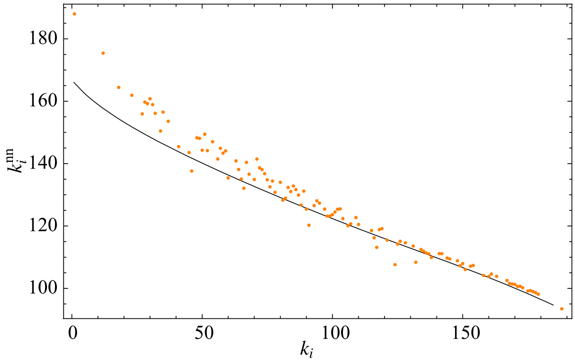


Fig. 7.8 Average nearest neighbour degree k_i^{nn} versus degree k_i in the undirected WTW (year 1995): (*points*) real data; (*solid line*) theoretical prediction of the fitness model

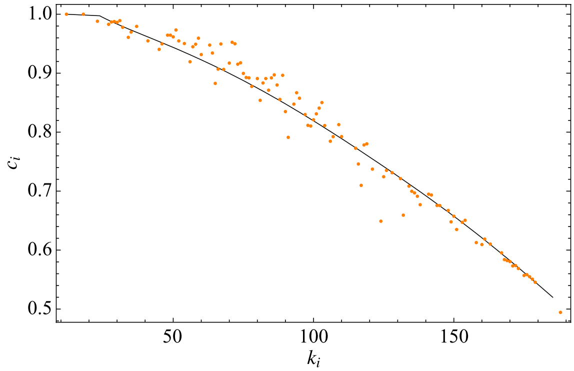


Fig. 7.9 Clustering coefficient c_i versus degree k_i in the undirected WTW (year 1995): (*points*) real data; (*solid line*) theoretical prediction of the fitness model

is suppressed. One can test the model prediction by calculating the expected average nearest neighbour degree as

$$\langle k_i^{\mathrm{nn}} \rangle = \frac{\sum_{j \neq i} \sum_{k \neq j} p_{ij} p_{jk}}{\sum_{j \neq i} p_{ij}}. \tag{7.16}$$

As Fig. 7.8 shows, the resulting curve represents again a very good prediction [43].

Finally, we report the behaviour of the clustering coefficient c_i, defined as the fraction of realised triangles originating at vertex i. This is shown in Fig. 7.9. Again, one finds a decreasing trend signalling that highly connected countries have on average poorly interconnected neighbours, while poorly connected countries have on average tightly interconnected neighbours. The model prediction can in this case be derived using the formula

$$\langle c_i \rangle = \frac{\sum_{j \neq i} \sum_{k \neq j,i} p_{ij} p_{jk} p_{ki}}{\sum_{j \neq i} \sum_{k \neq j,i} p_{ij} p_{ki}} \tag{7.17}$$

which, as shown in Fig. 7.9, is once again found to be in remarkable agreement with real data [43].

The above results suggest that the model reproduces the basic properties of the WTW, and traces them back to the heterogeneity in the GDPs of world countries.

Moreover, the simple assumptions at the basis of the model offer an interpretation for the observed properties of the WTW topology. In particular, the accordance between real data and the particular form of the model suggest that, once the heterogeneity at the level of individual countries is taken into account and assumed to directly affect the degrees $\{k_i\}$ alone, then most of the other network properties are indirectly and automatically explained [43]. An independent finding confirming this result comes from studies showing that the observed topology of the WTW is not significantly different from randomised variants obtained by keeping the original degree sequence fixed. Yet, the behaviour of higher-order properties such as k_i^{nn} and c_i, even if directly explained by that of the degrees, could not be simply predicted a priori without a quantitative model. This adds value to the modelling strategy presented here.

7.5.4 The Maximum Likelihood Principle

The results of the previous section are supported and refined by an inverse approach to the extraction of information from the WTW [39]. On a general ground, rather than assuming an empirical quantity as a candidate for the hidden variables $\{x_i\}$ and testing whether the networks generated with this choice are indeed similar to the real-world network considered, one can reverse the strategy and extract the values of the hidden variables $\{x_i\}$ directly from the real network. Then one can compare these unique values with candidate empirical quantities, to check whether a relation really exists. This comparison would automatically provide the form of the dependence between $\{x_i\}$ and the empirical quantities.

A way to realise this approach is provided by the Maximum Likelihood (ML) principle, a procedure commonly used in statistics whose use can be easily extended to networks [39]. Rather than requiring the values of model parameters as the input and generating an ensemble of possible networks as the output, the ML principle requires one particular real-world network as the input and provides the corresponding optimal parameter values $\{x_i^*\}$ as the output. *Optimal* stands for the parameter values that maximise the likelihood (or equivalently its logarithm) to obtain the particular real-world network under the model considered.

In the hidden variable model, the empirical input quantities are the rescaled GDP values $\{x_i\}$, which are fixed by observation, while the output values are the expected degrees $\{\langle k_i\rangle\}$. These expected values, and any other expected topological property, can then be compared with (but not fitted to) the empirical values $\{k_i\}$. By contrast, in the maximum likelihood approach the empirical input quantities are the degrees $\{k_i\}$, while the $\{x_i^*\}$ are output values depending uniquely on the observed degree sequence. In this case too, these output values can be compared with (but not fitted to) the empirical rescaled GDPs $\{x_i\}$ [39]. This comparison is shown in Fig. 7.10 where for consistency the parameter δ^* used in the hidden variable model and the same parameter used in the maximum likelihood approach have been both reabsorbed in the variables $\{x_i\}$ by redefining the latter as $x_i \to x_i\sqrt{\delta}$. One finds that the fitness values determined using only topological information are indeed proportional to the empir-

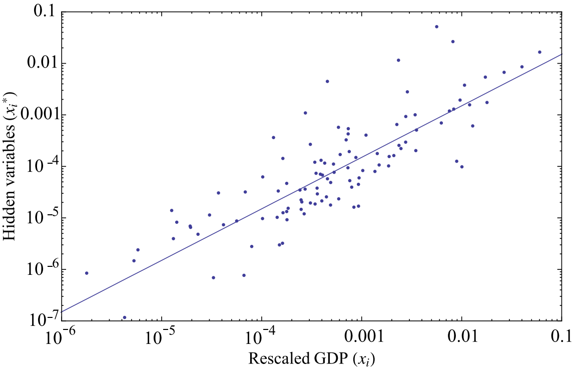


Fig. 7.10 Scatter plot of the likelihood-maximising values $\{x_i^*\}$ versus the rescaled GDP values $\{x_i\}$ (*isolated points*) and linear fit (*solid line*) for the undirected WTW in the year 2000

ical GDPs of world countries, and therefore that the maximum likelihood approach successfully identifies the GDP as the hidden variable shaping the topology of the WTW. Note that the two sets of values are, in principle, completely independent.

7.5.5 The WTW as a Directed Network

We now consider the topological properties of the WTW as a directed network. In any directed graph, the following relation exists between the entries $\{a_{ij}\}$ of its adjacency matrix and the entries $\{b_{ij}\}$ of the adjacency matrix of the same graph if regarded as undirected:

$$b_{ij} = a_{ij} + a_{ji} - a_{ij}a_{ji}. \tag{7.18}$$

Consequently, a relation is implied between the directed in- and out-degrees ($k_i^{\text{in}} = \sum_{j \neq i} a_{ji}$ and $k_i^{\text{out}} = \sum_{j \neq i} a_{ij}$) and the undirected degrees k_i observed on the two different representations of the same network:

$$k_i = k_i^{\text{in}} + k_i^{\text{out}} - k_i^{\leftrightarrow} \tag{7.19}$$

where $k_i^{\leftrightarrow} = \sum_{j \neq i} a_{ij}a_{ji}$ is the number of reciprocated connections incident at vertex i (the number of neighbours connected to i by incoming and outgoing links simultaneously), i.e. the reciprocated degree of vertex i.

The above relations indicate that, in principle, many different directed graphs can have the same undirected projection, making the latter not completely representative of an intrinsically directed network. However, the WTW has been found to display a peculiar structure that allows recovering significant information about its directed properties from the knowledge of the undirected ones [36, 37]. In particular, this is possible due to two empirically observed patterns. The first one is that on average $k_i^{\text{in}} \approx k_i^{\text{out}}$, i.e. a country generally has similar numbers of exporters and importers. The second one, shown in Fig. 7.11, is that the reciprocated degree $k_i^{\leftrightarrow}$ is proportional to the total degree $k_i^{\text{T}} \equiv k_i^{\text{in}} + k_i^{\text{out}}$

$$k_i^{\leftrightarrow} \approx \frac{r}{2} k_i^{\text{T}} \tag{7.20}$$

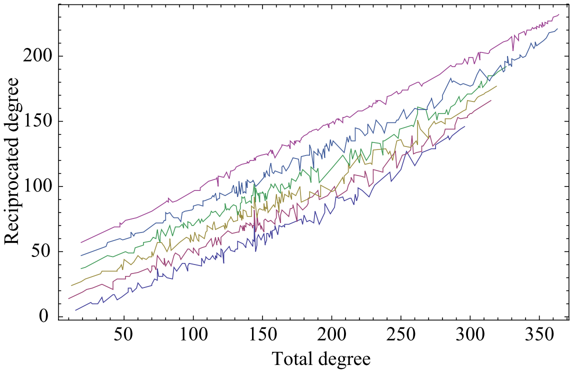


Fig. 7.11 Reciprocated degree $k_i^{\leftrightarrow}$ as a function of the total degree k_i^{T} in various directed snapshots of the WTW (from bottom to top: 1975, 1980, 1985, 1990, 1995, 2000)

where r is the *reciprocity* defined as the ratio between the number $L^{\leftrightarrow}$ of reciprocated links and the number L of directed links,

$$r \equiv \frac{L^{\leftrightarrow}}{L} = \frac{\sum_i \sum_{j \neq i} a_{ij} a_{ji}}{\sum_i \sum_{j \neq i} a_{ij}}. \tag{7.21}$$

The latter relation means that the number of simultaneous exporters and importers of a country is an approximately constant fraction of the total number of importers and exporters of the same country.

Combining the two observations mentioned above, it is possible to simply relate many directed properties of the WTW to its undirected ones [37]. For instance, the directed degrees can be expressed in terms of the undirected ones as

$$k_i^{\mathrm{in}} \approx k_i^{\mathrm{out}} \approx \frac{k_i^{\mathrm{T}}}{2} \approx \frac{k_i}{2-r}. \tag{7.22}$$

Similarly, the number L of directed links can be related to the number L^{u} of undirected links as follows:

$$L = \sum_i k_i^{\mathrm{in}} = \sum_i k_i^{\mathrm{out}} \approx \frac{\sum_i k_i}{2-r} = \frac{2L^{\mathrm{u}}}{2-r}. \tag{7.23}$$

Thus, the knowledge of the reciprocity r alone allows in many cases recovering the directed structure of the WTW from the undirected one. Note that this is a peculiar property of the WTW, as for a generic network no clear relation exists between the two representations, and a substantial loss of information may be associated with the undirected projection. Importantly, the above results hold for every analysed snaphot of the WTW [36] (see, for instance, Fig. 7.11). Therefore, the directedness of the network is easily monitored in terms of the time evolution of the reciprocity $r(t)$ [37]. The latter is shown in Fig. 7.12 together with a different measure of the reciprocity [36], i.e. the quantity

$$\rho \equiv \frac{r-c}{1-c}. \tag{7.24}$$

The index ρ is an alternative and more refined definition that allows consistent comparisons of the reciprocity across networks of different sizes and connectances [36].

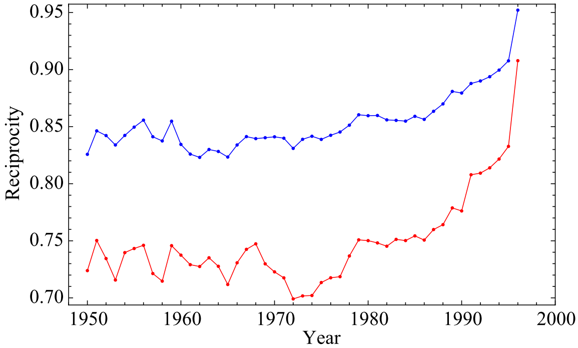


Fig. 7.12 Evolution of the reciprocity indices r (*top curve*) and ρ (*bottom curve*) in the directed WTW

Indeed, since the null value of r expected in the uncorrelated case (where there is no tendency towards either favouring or avoiding the formation of reciprocated links) is c itself, then r alone cannot be used to compare networks with different values of c. Thus, even if c displays small variations in time (as we showed in Fig. 7.4), these variations does not allow to assess the evolution of the reciprocity of the WTW on the basis of r alone. As Fig. 7.12 shows, both r and ρ display small fluctuations up to the early 1980s, and then increase rather steadily. However, the increase of ρ is steeper than that of r, signalling that (once density effects are taken into account) a rapid reciprocation process occurred in the WTW starting from the 1980s. It is instructive to combine this result with the approximately constant trend of the connectance in the undirected version of the WTW. As at the undirected level, there is no increase of link density, the rapid increase of reciprocity signals many new directed links being placed between countries that had already been trading in the opposite direction, rather than new pairs of reciprocal links being placed between previously non-interacting countries. In other words, many pairs of countries that had previously been trading only in a single direction have been establishing also a reverse trade channel, and this effect dominates on the formation of new bidirectional relationships between previously non-trading countries.

Since the directedness of the WTW can be significantly recovered from its undirected description in terms of the reciprocity parameter, and since the undirected network is excellently reproduced by the fitness model, a natural question is whether the model can be simply extended in order to reproduce the directed structure of the WTW. In the simplest case, the fitness model can be generalised to the directed case by introducing two fitness values $\{x_i, y_i\}$, separately controlling the out-degree and the in-degree of every vertex i. This allows to draw a directed link from i to j with an asymmetric probability $p_{ij} = p(x_i, y_j) \neq p_{ji}$. However, in this simple extension the presence of a link from i to j is statistically independent from the presence of the reciprocal link from j to i. This implies that, in contrast with what empirically observed, the reciprocity coefficient ρ is trivially zero, or in other words that there is no interesting reciprocity structure. Non-trivial reciprocity can only be generated by a more refined extension of the fitness model where mutual links are statistically dependent on each other. A way to do this is by drawing, for each single vertex pair (i, j), a single link from i to j, a single link from j to i, two reciprocal links, or no link at all with four different probabilities (that must sum up to one) [36, 38]. In this generalised model, each vertex is now assigned three fitness values $\{x_i, y_i, w_i\}$, separately controlling its non-reciprocated outgoing links, non-reciprocated incoming links, and reciprocated links going both ways, respectively [38]. When applied to the directed WTW, this model turns out to reproduce all the properties discussed above. Notably, the model preserves a high degree of simplicity, as the three values $\{x_i, y_i, w_i\}$ are all again related to the GDP alone.

These results highlight once again that the GDP is the main factor underlying the unweighted topology of the WTW, and that a satisfactory and detailed model of the network can be defined in terms of the empirical GDP values. However, as we mentioned, when weights are explicitly considered a range of new possibilities emerge, as relations that are indistinguishable at a topological level may be strongly heterogeneous at the weighted level. As we discussed, while nonzero weights are well reproduced by gravity models of trade, there is currently no model that allows to simultaneously capture the topology and the weighted architecture of the real WTW. This implies that one important open problem to address in the future is the definition of a unified framework where network models and gravity models are reconciled. This is likely to involve joint efforts from the different scientific communities of trade economists and network theorists.

References

1. Almeida, H., Wolfenzon, D.: A theory of pyramidal ownership and family business groups. J. Finance **61**(6), 2637–2680 (2006)
2. Anderberg, M.R.: Cluster Analysis for Applications. Academic Press, New York (1973)
3. Aste, T., Di Matteo, T., Hyde, S.: Complex networks on hyperbolic surfaces. Physica A **346**, 20–26 (2005)
4. Banzhaf, J.F.: Weighted voting doesn't work: A mathematical analysis. Rutgers Law Rev. **19**, 317–343 (1965)
5. Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A.: The architecture of complex weighted networks. Proc. Natl. Acad. Sci. USA **101**, 3747 (2004)
6. Battiston, S., Catanzaro, M.: Statistical properties of corporate board and director networks. Eur. Phys. J. B **38**(2), 345–352 (2004)
7. Battiston, S., Catanzaro, M.: Statistical properties of corporate board and director networks. Eur. Phys. J. B **38**, 345–352 (2004)
8. Battiston, S., Rodrigues, J.F., Zeytinoglu, H.: The network of inter-regional direct investment stocks across Europe. Adv. Complex Syst. **10**(1), 29–51 (2005)
9. Bonanno, G., Vandewalle, N., Mantegna, R.: Taxonomy of stock market indices. Phys. Rev. E **62**, 7615–7618 (2000)
10. Boss, M., Elsinger, H., Summer, M., Thurner, S.: Network topology of the interbank market. Quant. Finance **4**, 677–684 (2004)
11. Brioschi, F., Buzzacchi, L., Colombo, M.: Risk capital financing and the separation of ownership and control in business groups. J. Bank. Finance **13**(1), 747–772 (1989)
12. Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., Wiener, J.: Graph structure of the web. Comput. Netw. **33**, 309–320 (2000)
13. Caldarelli, G.: Scale-Free Networks: Complex Webs in Nature and Technology. Oxford University Press, Oxford (2007)

14. Caldarelli, G., Capocci, A., De Los Rios, P., Munoz, M.A.: Scale-free networks from varying vertex intrinsic fitness. Phys. Rev. Lett. **89**, 258702 (2002)
15. Chapelle, A.: Separation of ownership and control: Where do we stand? Corp. Ownersh. Control **15**(2) 91–101 (2005)
16. Chapelle, A., Szafarz, A.: Controlling firms through the majority voting rule. Physica A **355**, 509–529 (2005)
17. Coronnello, C., Tumminello, M., Lillo, F., Micciche, S., Mantegna, R.: Sector identification in a set of stock return time series traded at the London stock exchange. Acta Phys. Pol. B **36**, 2653–2679 (2005)
18. Corrado, R., Zollo, M.: Small worlds evolving: Governance reforms, privatizations, and ownership networks in Italy. Ind. Corp. Change **15**(2), 319–352 (2006)
19. Cubbin, J., Leech, D.: The effect of shareholding dispersion on the degree of control in British companies: Theory and measurement. Econ. J. **93**(370), 351–369 (1983)
20. Davis, G.F., Yoo, M., Baker, W.E.: The small world of the corporate elite, 1982–2001. Strateg. Organ. **1**, 301–326 (2003)
21. De Bandt, O., Hartmann, P.: Systemic risk: a survey. CEPR Discussion Paper No. 2634 (2000)
22. De Masi, G., Gallegati, M.: Debt-credit economic networks of banks and firms: The Italian case. In: Chatterjee, A., Chakrabarti, B.K. (eds.) Econophysics of Markets and Business Networks. Springer, Berlin (2007)
23. De Masi, G., Iori, G., Caldarelli, G.: Fitness model for the interbank money market. Phys. Rev. E **74**, 066112 (2006)
24. De Masi, G., Fujiwara, Y., Gallegati, M., Greenwald, B., Stiglitz, J.: An analysis of the Japanese credit network (2009). arxiv:0901.2384
25. Derrida, B., Flyvbjerg, H.: Multivalley structure in Kauffman's model: Analogy with spin glasses. J. Phys. A **19**, 1003–1008 (1986)
26. Di Matteo, T., Aste, T., Hyde, S., Ramsden, S.: Interest rates hierarchical structure. Physica A **335**, 21–33 (2004)
27. Dore, R.: Stock market capitalism vs. welfare capitalism. New Polit. Econ. **7**(1), 115–127 (2002)
28. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of Networks: From Biological Nets to the Internet and WWW. Oxford University Press, Oxford (2003)
29. Eisenhardt, K.M.: Agency theory: An assessment and review. Acad. Manag. Rev. **14**(1), 57–74 (1989)
30. Lillo, F., Moro, E., Vaglica, G., Mantegna, R.N.: Specialization and herding behavior of trading firms in a financial market. New J. Phys. **10**, 043019 (2008)
31. Fagiolo, G., Reyes, J., Schiavo, S.: World-trade web: Topological properties, dynamics, and evolution. Phys. Rev. E **79**, 036115 (2009)
32. Fagiolo, G., Reyes, J., Schiavo, S.: On the topological properties of the world trade web: A weighted network analysis. Physica A **387**, 3868–3873 (2008)
33. Feenstra, R.C., Yang, T.H., Hamilton, G.G.: Business groups and product variety in trade: evidence from South Korea, Taiwan and Japan. J. Int. Econ. **48**(1), 71–100 (1999)
34. Flath, D.: Indirect shareholding within Japan's business groups. Econ. Lett. **38**(2), 223–227 (1992)
35. Fujiwara, Y., Aoyama, H., Ikeda, Y., Iyetomi, H., Souma, W.: Structure and temporal change of the credit network between banks and large firms in Japan. Economics **3**, 2009-7 (2009)
36. Garlaschelli, D., Loffredo, M.: Fitness-dependent topological properties of the world trade web. Phys. Rev. Lett. **93**, 268701 (2004)
37. Garlaschelli, D., Loffredo, M.: Structure and evolution of the world trade network. Physica A **355**, 138–144 (2004)
38. Garlaschelli, D., Loffredo, M.: Multispecies grand-canonical models for networks with reciprocity. Phys. Rev. E **73**, 015101 (2006)
39. Garlaschelli, D., Loffredo, M.: Maximum likelihood: extracting unbiased information from complex networks. Phys. Rev. E **78**, 015101 (2008)
40. Garlaschelli, D., Caldarelli, G., Pietronero, L.: Universal scaling relations in food webs. Nature **423**(6936), 165–168 (2003)

41. Garlaschelli, D., Battiston, S., Castri, M., Servedio, V.D.P., Caldarelli, G.: Detecting communities in large networks. Physica A **350**, 491–499 (2005)
42. Garlaschelli, D., Di Matteo, T., Aste, T., Caldarelli, G., Loffredo, M.: Interplay between topology and dynamics in the world trade web. Eur. Phys. J. B **57**, 159 (2007)
43. Garlaschelli, D., Loffredo, M.I.: Fitness-dependent topological properties of the world trade web. Phys. Rev. Lett. **93**(18), 188701 (2004). doi:10.1103/PhysRevLett.93.188701
44. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. Proc. Natl. Acad. Sci. USA **99**, 7821–7826 (2002)
45. Glattfelder, J.B., Battiston, S.: Backbone of complex networks of corporations: The flow of control. Phys. Rev. E **80**(3), 36104 (2009)
46. Gleditsch, K.: Expanded trade and GDP data. J. Confl. Resolut. **46**, 712–724 (2002)
47. Gopikrishnan, P., Rosenow, B., Plerou, V., Stanley, H.: Quantifying and interpreting collective behavior in financial markets. Phys. Rev. E **64**, 035106 (2001)
48. Gower, J.: Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika **53**, 325–338 (1966)
49. Granovetter, M.: Coase revisited: Business groups in the modern economy. Ind. Corp. Change **4**, 93–130 (1995)
50. Herfindahl, O.: Copper Costs and Prices: 1870–1957. John Hopkins University Press, Baltimore (1959)
51. Iori, G., Jafarey, S., Padilla, F.: Systemic risk on the interbank market. J. Econ. Behav. Organ. **61**, 525–542 (2006)
52. Iori, G., Renó, R., De Masi, G., Caldarelli, G.: Trading strategies in the Italian interbank market. Physica A **376**, 467–479 (2007)
53. Iori, G., De Masi, G., Precup, O., Gabbi, G., Caldarelli, G.: A network analysis of the Italian overnight money market. J. Econ. Dyn. Control **32**, 259–278 (2008)
54. Kogut, B., Walker, G.: The small world of Germany and the durability of national networks. Am. Sociol. Rev. **66**(3), 317–335 (2001)
55. La Porta, R., de Silanes, F.L., Shleifer, A.: Corporate ownership around the world. J. Finance **54**(2) 471–517 (1999)
56. La Porta, R., de Silanes, F.L., Shleifer, A., Vishny, R.: Law and finance. J. Polit. Econ. **106**(6), 1113–1155 (1998)
57. Leech, D.: The relationship between shareholding concentration and shareholder voting power in British companies: A study of the application of power indices for simple games. Manag. Sci. **34**(4), 509–527 (1988)
58. Mantegna, R.: Hierarchical structure in financial markets. Eur. Phys. J. B **11**, 193–197 (1999)
59. Mantegna, R., Stanley, H.: An Introduction to Econophysics: Correlations and Complexity in Finance. Cambridge University Press, Cambridge (2000)
60. Miccichè, S., Bonanno, G., Lillo, F., Mantegna, R.: Volatility in financial markets: stochastic models and empirical results. Physica A **314**, 756–761 (2003)
61. Miceli, M., Susinno, G.: Using trees to grow money. Risk 11–12 (November 2003)
62. Newman, M., Strogatz, S., Watts, D.: Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E **64**(2), 26118 (2001)
63. Newman, M., Watts, D., Strogatz, S.: Random graph models of social networks. Proc. Natl. Acad. Sci. USA **99**, 2566–2572 (2002)
64. Onnela, J.P., Kaski, K., Kertesz, J.: Clustering and information in correlation based financial networks. Eur. Phys. J. B **38**, 353–362 (2004)
65. Onnela, J.P., Chakraborti, A., Kaski, K., Kertesz, J., Kanto, A.: Asset trees and asset graphs in financial markets. Phys. Scr. T **106**, 48–54 (2003)
66. Onnela, J.P., Saramäki, J., Kertész, J., Kaski, K.: Intensity and coherence of motifs in weighted complex networks. Phys. Rev. E **71**(6), 065103 (2005)
67. Park, J., Newman, M.: Properties of highly clustered networks. Phys. Rev. E **68**, 026112 (2003)
68. Pastor-Satorras, R., Vázquez, A., Vespignani, A.: Dynamical and correlation properties of the Internet. Phys. Rev. Lett. **87**(25), 258701 (2001)

69. Serrano, M.: Phase transition in the globalisation of trade. J. Stat. Mech. Theory Exp. (2007). doi:10.1088/1742-5468/2007/01/L01002
70. Serrano, M., Boguñá, M.: Topology of the world trade web. Phys. Rev. E **68**(1), 15101 (2003)
71. Serrano, M.A., Boguñá, M., Vespignani, A.: Patterns of dominant flows in the world trade web. J. Econ. Interact. Coord. **2**, 111–124 (2007)
72. Shapley, L.S., Shubik, M.: A method for evaluating the distribution of power in a committee system. Am. Polit. Sci. Rev. **48**(3), 787–792 (1954)
73. Shleifer, A., Vishny, R.W.: A survey of corporate governance. J. Finance **52**(2), 737–783 (1998)
74. Sorämaki, K., Bech, M., Arnold, J., Glass, R., Beyeler, W.: The topology of the interbank payment flows. Physica A **379**, 317–333 (2007)
75. Tinbergen, J.: Shaping the World Economy: Suggestions for an International Economic Policy. The Twentieth Century Fund, New York (1962)
76. Tumminello, M., Lillo, F., Mantegna, R.N.: Correlation, hierarchies, and networks in financial markets. Journal of Economic Behavior and Organization **75**, 40–58 (2010)
77. Tumminello, M., Aste, T., Di Matteo, T., Mantegna, R.: A tool for filtering information in complex systems. Proc. Natl. Acad. Sci. USA **102**, 10421–10426 (2005)
78. Tumminello, M., Coronnello, C., Lillo, F., Miccichè, S., Mantegna, R.: Spanning trees and bootstrap reliability estimation in correlation-based networks. Int. J. Bifurc. Chaos Appl. Sci. Eng. **17**, 2319–2329 (2007)
79. Vitali, S., Glattfelder, J.B., Battiston, S.: The map of the global corporate control network. Working paper (2008)