J Econ Interact Coord (2010) 5:191–219
DOI 10.1007/s11403-010-0068-4

REGULAR ARTICLE

Ownership and control in shareholding networks

Giulia Rotundo · Anna Maria D'Arcangelis

Received: 27 August 2009 / Accepted: 23 June 2010 / Published online: 10 July 2010

Abstract This paper aims to provide a network analysis of the relationships of shareholders in the Italian stock market for understanding the relevance of portfolio diversification in integrated ownership and firms control. The analysis combines both a complex network and an operational research approach. The former is used for statistical analyses on portfolio diversification. The latter estimates integrated ownership considering the paths on the network, and it emphasizes the difference between ownership and control. The dataset consists of nearly 300 companies traded on the Italian Stock Market. Data were retrieved through CONSOB and AIDA database, and they are adjourned at 2008. The dataset is completed with information on banks and insurance companies. Such data were retrieved through BANKSCOPE and ISIS databases.

Keywords Stock market · Portfolio diversification · Shareholding · Ownership · Control

1 Introduction

Modelization through networks has been shown to be a powerful instrument for understanding the behaviour of complex systems, and network models of financial data have already pointed out many relevant properties of stock markets. The present report uses a network representation in order to perform an analysis of the structure of ownership of firms listed in the Italian Stock Market.

G. Rotundo (✉) · A. M. D'Arcangelis
Faculty of Economics, University of Tuscia, Via del Paradiso 47, 01100 Viterbo, Italy
e-mail: Giulia.Rotundo@uniroma1.it

A. M. D'Arcangelis
e-mail: annamaria.darcangelis@gmail.com

We propose the new target to understand to which extent the diversification of shareholdings in company's portfolios gives a good estimate of the relevance of the company in the market with respect to ownership and control of other companies. This, in turn, allows to determine whether portfolio diversification is a common strategy for acquiring control, to understand the weight in the market of financial companies like as banks and insurance firms, and ultimately to detect oligopolies and to add issues to the analysis of risk.

The data sets we analyze, adjourned at the May 10th 2008, report information on shareholders and subsidiaries of companies traded on the MTA segment of the Italian Stock Market. The active and passive ownership sample is build extracting records from Bureau Van Dijk databases: Bankscope for banking and financial companies; ISIS for insurance companies; AIDA for all the remaining sectors. Furthermore, the shareholdings data (passive holdings) are checked with the "Ownership data" published by CONSOB, based on the notifications submitted by persons holding, directly or indirectly, more than 2% of share capital, in accordance with the prevision of article 120 of Legislative Decree 58 of 24 February 1998[1]. Balance-sheet data on equity capitalizations and total asset values are extracted from Datastream Thomson Financial Database.

The unavailability of any of the information needed (active or passive holdings) produced the exclusion of the company from the sample. The availability of several data providers permitted us to consider also very limited holdings (below 2%); yet we exclude all the mediate possessions held via mutual funds.

The network analysis is limited to holdings among listed firms (active or passive ownership of listed companies in listed companies). The resulting size of the sample is 247, i.e. 94% of the universe of listed companies and 95.22% in terms of capitalization.

The dataset differs from the one examined in Garlaschelli et al. (2005) because the bundle of listed companies in 2008 is slightly different from 2002 (and for the accuracy due to the level of detail of ownership data).

We start our analysis building a network from data. Each company corresponds to a node and a link from node i to node j exists if i owns shares of j. Therefore, we obtain a directed graph and the direction of actual links is the opposite of the ones of Garlaschelli et al. (2005), but the same used in Chapelle and Szafarz (2005), Salvemini et al. (1995). In our network construction, the number of links exiting from a node, i.e. the out-degree k_{out}, is a quantitative measure of portfolio diversification. In line with a remark of Dietzenbacher and Temurshoev (2008), and remembering the reversal direction of the links, we also considered the number of links entering in a node k_{in}, that shows the number of listed companies owning shares of each firm in the sample. Both k_{in} and k_{out} are biased due to the sampling method, involving only listed companies. Nevertheless, the degree of sampling bias is more severe for k_{in}, given the fundamental role of this figure in the evaluation of ownership structure. Therefore, we proceed to analyze k_{out} only.

We proceed evaluating portfolio diversification, wealth, ownership, control, and the correlations among them (Chapelle and Szafarz 2005; Salvemini et al. 1995).

[1] For more details see http://www.consob.it/main/emittenti/societa_quotate/index.html.

Portfolio diversification is measured through mere statistical analysis of the probability distribution of k_{out}, through the estimate of the assortativity (Newman 2003; Newman et al. 2006) and through the analysis of the paths that are hierarchical in accord with the measure introduced in Trusina et al. (2004).

Such hierarchy measure quantifies the percentage of hierarchical paths connecting pairs of nodes. Two nodes i and j are in a hierarchical path if an "up" path exists from node i through nodes with higher k_{out}, followed by a "down" path where nodes on the path have a decreasing k_{out}. Therefore, a hierarchical path connects nodes i and j if they have a common ancestor, i.e. if they are in the portfolio of a larger investor. The number of hierarchical paths measures the relevance of portfolio diversification. The fraction of paths that are hierarchical is an information different from the mere node rank, and it is positively correlated to degree correlation and assortativity (Trusina et al. 2004).

We also compare the above results with the market value of the portfolio. This allows to measure the overlap between portfolio diversification and company size, and to answer to questions on the possibility to diversify portfolios for companies having a smaller capitalization.

It is worth remarking that the analysis of ownership/control is based on input-output matrix approach for indirect ownership computation (Chapelle and Szafarz 2005; Dietzenbacher and Temurshoev 2008; Huber and Ryll 1989; Salvemini et al. 1995; Turnovec 1999, 2005). In fact, while direct ownership in a sector can be readily obtained from data on shareholdings, the true ownership structure may be hidden by a complex network of indirect relations. Therefore, a further analysis of the ownership matrix analysis is needed, and we switch from the realm of Complex Networks to the one of Operations Research. While input-output methods were proposed in the literature since 1989 by Huber and Ryll, we had to select the method most proper to our data. In fact, networks without loops are considered in Salvemini et al. (1995); and a neat separation among individuals and companies is used in Turnovec (1999), Turnovec (2005), while we consider companies only. Moreover, in Turnovec (1999), Turnovec (2005) all the shares are owned by individuals and companies listed in the sample, whilst this condition does not hold for the sample of companies traded on one selected Stock Markets. We follow the approach of Chapelle and Szafarz (2005) because, under mild hypotheses, it is well suitable for networks with cycles and it provides a clear analysis of the role of each company in the system avoiding double counting problems not considered in the previous papers. We refer to Chapelle and Szafarz (2005) also for the definition of control, leaving the analysis of other means of control for future work.

The paper is organized as follows: the first step is to examine the network structure, paying attention to connected components and cycles. Then, we perform statistical analyses on portfolio diversification estimating the probability distribution of k_{out}. Assortativity and hierarchicity complete the analysis of the overall structure of the network giving information of node degree correlation. Then we consider capitalization and issues related to wealth invested. Section 4 deals with the ownership and control structure. Section 5 adds a remark on the giant component. Finally, Sect. 6 gives the conclusion.

2 Network structure: connected components and cycles

The detection of connected components and cycles is relevant for the possibility to apply some algorithms for the detection of control, therefore we start examining connected components and cycles.

The shareholding network has 247 nodes, and 243 links, and each node is a company traded on the Stock Market. We can look for cross-ownership into connected components, therefore we first detect weakly connected components. There are 130 isolated nodes, one giant weakly connected component (101 nodes, and 187 links), and 7 smaller size weakly connected components (5 of them have 2 nodes only, and the remaining 2 have 3 nodes). Figure 1 shows the giant component, and Fig. 2 shows the other weakly connected components, apart from isolated nodes and self-loops. Table 1 reports the names of the isolated nodes, i.e. the companies that neither own shares of other companies traded on the Italian Stock Market, nor are owned themselves by other companies.

The absence of cycles is relevant for the application of algorithms for paths (Salvemini et al. 1995). Cycles and their relevance is evidenced by the analysis of the strongly connected components: the ones which cardinality is bigger than one are 4 only, and they prevent the usage of algorithms for directed acyclic graphs. Table 2 lists the strongly connected components having 2–3 nodes each. The largest strongly connected component contains 12 nodes, and 34 links, and it is shown in Fig. 3. The entire network becomes acyclic only whether links corresponding to more 10% share ownerships are considered. Hence, cycles cannot be eliminated without cutting

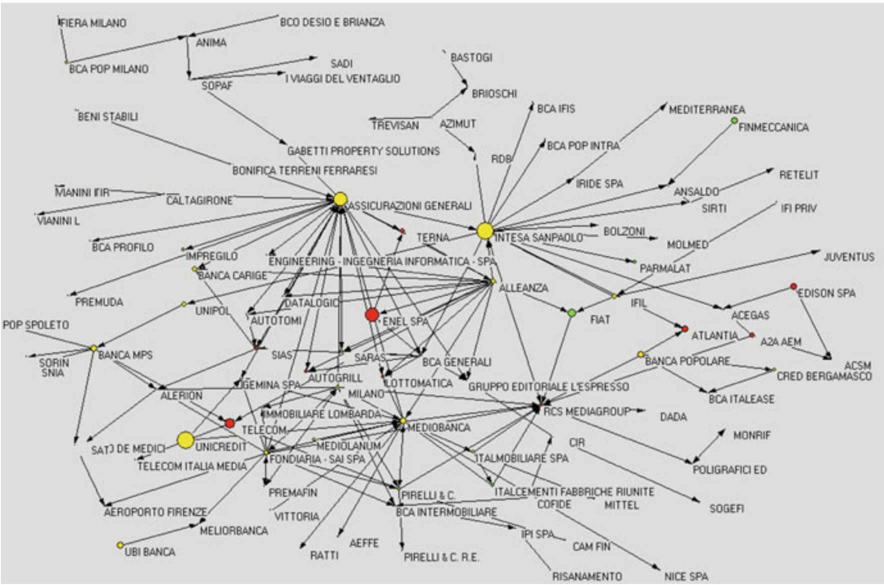

Fig. 1 The giant weak connected component of the shareholding network showing financial (*white*), industrial (*light grey*), and services (*dark grey*) nodes. Nodes size is proportional to the capitalization of the company represented by the node

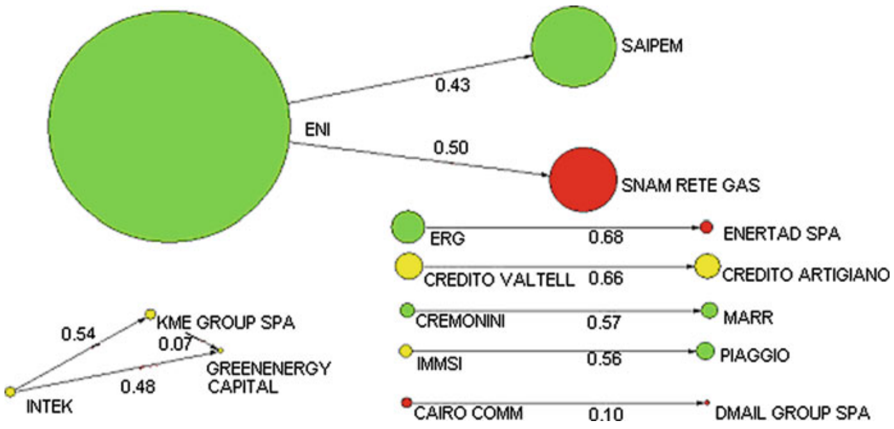

Fig. 2 Small weak connected components in the shareholding network showing financial (*white*), industrial (*light grey*), and services (*dark grey*) nodes. Nodes size is proportional to the capitalization of the company represented by the node

relevant information, and we need to look for other methods for the estimate of integrated ownership and control.

The sources in this networks are the companies that are not owned by other companies. There are 16 sources, listed in Table 3. These sources are companies themselves, so we do not follow the approach of Turnovec (1999, 2005) separating the sources by the other nodes, since in our sample sources are not individuals. Moreover, since the shares are not totally owned and sold inside this market, a further hypothesis used in Turnovec (1999, 2005) does not hold.

The topology of the network suggests that the small amount of 16 sources might control both the weakly and the strongly connected components. However, further inquiry is needed to check this hypothesis, and to understand whether the control is achieved through either long or small chains, and the role that portfolio diversification is playing into this scheme.

3 Portfolio diversification

In our network each node is a company. Let E be the set of nodes. A link from node i to node j exists if i owns shares of $j, \forall i, j \in E$. Therefore, the number of links exiting from a node measures portfolio diversification. Given the set of nodes E, we build the vector $k = (k^i_{\text{out}})$, in which each component, k^i_{out}, represents the number of links exiting from node i. The values k^i_{out} can also be considered as the values that the variable k_{out} assumes on node i. We start analyzing the probability distribution of k_{out}, and we compare the results with the one reported in Garlaschelli et al. (2005). Then, we add insights on the node degree correlation through the estimate of assortativity and hierarchy level of the network Trusina et al. (2004).

Table 1 0-connected companies

Financials	Industrials	Industrials	Services
Aedes	Danieli	Mirato	CAD IT
Banca Finnat	Davide Campari SPA	Montefibre	Caltagirone ED
Bca Pop Etruria	De Longhi	Negri Bossi	CDC
Bco Sardegna	Diasorin SPA	Olidata	Chl
Cattolica Ass	Ducati	Panaria Group	Class
Credem	Eems Italia Spa	Permasteelisa	Dataservice
Ergo Previdenza Spa	El.En. Spa	Pininfarina	Digital Bros
Finarte Casa D'aste	Elica SPA	Poligrafica SF	Dmt Didital Tech
Gruppo Mutuionlin	Emak	Poltrona Frau	Enia SpA
IGD	Eurotech	Prima Ind	Esprinet
Investimenti e Sviluppo SPA	Everel Group	Prysmian	Eurofly
Snai	Fidia	Recordati	Eutelia
Industrials	Filatura di Pollone	Richard Ginori	Exprivia
Aicon	Gefran	Sabaf	Fastweb
Amplifon	Geox	Saes Getters	FMR-Art'e'
Arena	Gewiss	Safilo	Fullsix
Astaldi	Granitifiandre	Schiapparelli	Gas Plus
Basicnet	Gruppo Ceramiche Ricchetti	Socotherm	Gruppo Coin
Beghelli	Guala Closures	Sol	Hera
Benetton Group	Ima	Stefanel	I Grandi Viaggi
Bialetti	Indesit	Tods	Il Sole 24 ore
Biesse	Interpump SPA	Trevi	Itway
Boero Bartolomeo	Irce SPA	Zucchi	Kaitech
Brembo	Isagro	Zignago	Mediaset
Bulgari	IT holding	*Services*	Mondo tv
Buzzi Unicem	La Doria	Acea	Navigazione Montanari
Carraro	Lavorwash	Acotel Group SPA	Omnia Network
Cembre	Luxottica	Acque Potabili	Reply
Cent Latte Torino	Maffei	Actelios	Seat p.g.
Cobra	Maire Tecnimont	Alitalia	Tas
Crespi	Marazzi	ASRoma	Tiscali
Csp	Marcolin	Ascopiave	Txt e sol
Damiani	Mariella Burani Fashion	Buongiorno	Mondadori

Number of isolated elements in the macro-sectors: financial 12; industrial 77; services 41

3.1 Probability distribution of k_{out}

The present section reports the analysis of nodes degree k_{out}. The approach is the same used in Garlaschelli et al. (2005), so a direct comparison of results is possible. We estimate both the probability distribution $P(k_{\text{out}})$ of the number of vertices with in-degree greater than or equal to k_{out} and the probability density $p(k_{\text{out}})$ for

Table 2 The three small-size strongly connected components

Owner	Owned
Alerion	Banca MPS (7.96)
Banca MPS	Alerion (7.06)
Monrif	Poligrafici ed (60.02)
Poligrafici ed.	Monrif (0.08)
Fondiaria–SAI SPA	Milano (51.098)
Fondiaria–SAI SPA	Premafin (4.47)
Milano	Fondiaria SAI–SPA (6.69)
Premafin	Fondiaria SAI–SPA (37.2720)
Premafin	Milano (0.3)
Milano	Premafin (2.23)

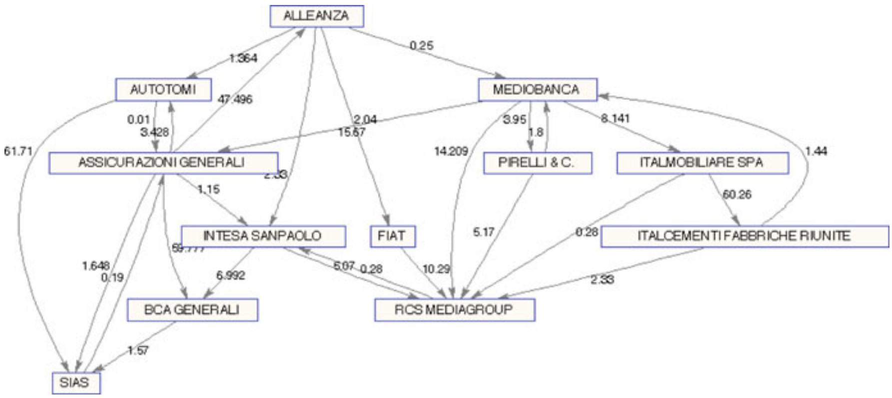

Fig. 3 The largest strongly connected component (12 nodes, and 34 links). Self-loops are not drawn

companies that own at least another company. We test the hypothesis of power law curves. In order to validate the hypothesis, both the conditions $P(k_{\text{out}}) \sim k_{\text{out}}^{1-\gamma}$ and $p(k_{\text{out}}) \sim k_{\text{out}}^{-\gamma}$ must be validated. The best fit of the probability distribution gives $\gamma = 1.7925$ (1.6596, 1.9254), the best fit of the density gives $\gamma = 2.159$ (1.984, 2.339) (Fig. 4), and the estimate of γ though the Maximum Likelihood Estimate (MLE) gives $\gamma = 2.72766$ (2.72763, 2.72768).

The confidence intervals do not overlap. The Jarque–Bera test on the normality of residuals accepts the power law for the density case, but rejects it for the distribution case. Therefore, we can state that portfolio diversification k_{out} does not follow a power law. Figure 4 shows the histograms and the regression curves.

Table 3 Sources belonging to the giant connected component

A2a Aem	BCO Desio e Brianza	Cofide	Risanamento
Banca Popolare	Bonifica Terreni Ferraresi	Edison SPA	UBI Banca
Bastogi	Caltagirone	Finmeccanica	Unicredit
BCA Pop Milano	Cam Fin	Mediolanum	Vittoria

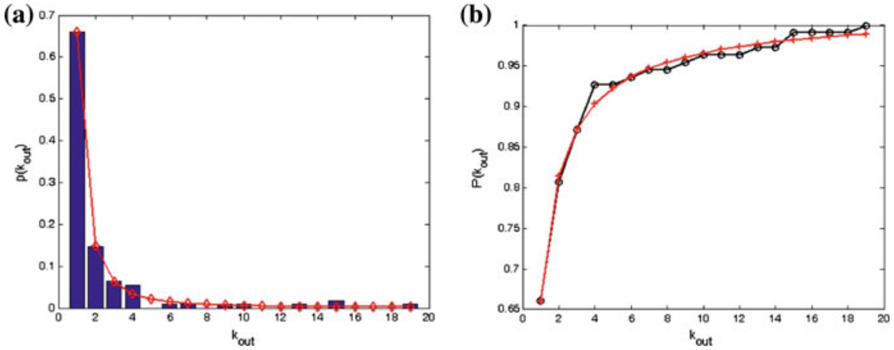

Fig. 4 Entire dataset. **a** Histogram (*bars*) and power law best fit (continuous *grey curve* with diamonds). Estimate of $p(k_{\text{out}}) \sim k_{\text{out}}^{-\gamma}$ gives $\gamma = 2.159(1.984, 2.339)$, RMSE = 0.0094. The Jarque–Bera test validates the hypothesis of Gaussianity of residuals. **b** Empirical probability distribution (*black line* and *circles*) and power law best fit (continuous *grey curve* with diamonds). Estimate of $P(k_{\text{out}}) \sim k_{\text{out}}^{1-\gamma}$ gives $\gamma = 1.7925\,(1.6596, 1.9254)$, RMSE = 0.0088. The MLE γ gives $\gamma = 2.72766\,(2.72763, 2.72768)$. The Jarque–Bera test rejects the hypothesis of Gaussianity of residuals

A comparison can be performed among our Figs. 4 and 2 of Garlaschelli et al. (2005) (inset). The comparison is related to different data sets, as the companies trading on the stock market in 2002 are different form the ones of 2008. However a global behavior of the companies can be evidenced. We remark that the decay that we detected is sharper than the decay pointed out in Garlaschelli et al. (2005), therefore the trend in diversification has been to lower diversification, while the maximum of diversification remained fixed at 19. Without considering the 0-connected companies, the mean of k_{out} is 2.14. The mode and the median are equal to 1.

Companies in our dataset are grouped into three areas:[2] financial ($n = 65$), industrial ($n = 109$), and services ($n = 73$). We repeat the analysis considering these macro-sectors. Table 4 reports the list of the most connected companies in each sector. For the sake of completeness, we also report the list of 1-connected components (Table 5). In the financial sector the mean of k_{out} is 3.15; in the industrial sector the mean of k_{out} is 1.37; in the services sector the mean of k_{out} is 1.25. The mode and the median are always equal to 1.

It appears that the most connected companies belong to the financial sector, the result is in line with the role of financial companies. Figure 5 shows the analysis of k_{out} for the financial sector. A power law is detected with exponent $\gamma \in (2.22, 2.36)$. The MLE gives $\gamma = 2.64\,(2.35, 2.94)$, so leading to $\gamma \cong 2.35$. The low level of diversification of the industrial and services sectors does not allow to repeat the regression analysis for the estimate of $P(k_{\text{out}}) \sim k_{\text{out}}^{1-\gamma}$ and $p(k_{\text{out}}) \sim k_{\text{out}}^{-\gamma}$ within a statistically significant level.

From the analyses performed above we can conclude that portfolio diversification is performed by a small minority of companies, while most of them own at most just another company. The fact that financial companies own more companies can be easily

[2] We follow the classification scheme that has been proposed by Borsa Italiana S.p.A. See http://www.borsaitaliana.it/bitApp/curiosita.bit?target=Settori.

Table 4 The table reports the most connected companies in the three macro sector (financials, industrials, services) of the Italian Stock Market

Financials	k_{out}	Industrials	k_{out}	Services	k_{out}
Assicurazioni Generali	19	Pirelli & C.	4	A2A	3
Alleanza Ass.ni	15	Eni	3	RCS Mediagroup	3
Intesa San Paolo	15	Italcementi	3	Sias	2
Fondiaria Sai	13	Bonifiche Ferraresi	2	Autostrada To-Mi	2
Milano Assicurazioni	10	Cremonini	2	Edison spa	2
Mediobanca	9	Erg	2	Poligrafici Editoriale	2
Banca Generali	7	Fiat	2		
Banca Mps	6	Sirti	2		
Azimut	4				
Banca Popolare	4				
Caltagirone	4				
Ifil	4				
Italmobiliare	4				
Cir	3				
Sopaf	3				
Ifi Priv	3				
Alerion	2				
Banca Intermobiliare	2				
Banca Popol. Milano	2				
Cofide	2				
Intek	2				
Premafin	2				
Unicredit	2				

Companies that are not listed in the table either own at most the shares of only one company or are owned by at most another company. Without considering the 0-connected companies, the mean of k_{out} is 2.14. In the financial sector the mean of k_{out} is 3.15; in the industrial sector the mean of k_{out} is 1.37; in the services sector the mean of k_{out} is 1.25. The mode and the mean are always equal to 1

explained through their institutional role, and it is supported by a recent evidence of the evolution of the ownership and control structures in Italy (Bianchi et al. 2008), pointing out the increasing role of institutional investors in domestic companies' ownership between 1990 and 2007.[3] So we note that portfolio diversification seems to be due to practical managerial issues more than to instances of expanding on the market, and that this attitude shrank if comparing to 2002 data. Further analyses are needed to validate this hypothesis.

[3] The paper, jointly authored by Bank of Italy and Consob researchers, also note that the limited weight of banks in the capital of Italian companies did not change in the period: "whereas institutional investors are present approximately in the same percentage of banks over the whole period, in 2007 they own stakes in twice the number of non financial companies than in 1990. So the result is entirely due to the increasingly presence of other financial intermediares. Aganin and Volpin (2005), Sect. 2.3.

Table 5 1-connected companies

Financials	Industrials	Services
Aedes	Amplifon	Acotel Group
Anima	Boero Bartolomeo	Ascopiave
Banca Finnat	Buzzi Unicem	Buongiorno
Bastogi	Damiani	Cairo Comm
Bca Ifis	Danieli	Caltagirone Ed
Bco Desio e Brianza	Datalogic	Cam Fin
Beni Stabili	Davide Campari	Cdc
Cred Bergamasco	Elica	Enel
Igd	Fidia	Esprinet
Immsi	Finmeccanica	Exprivia
Mediolanum	Gefran	Fiera Milano
Monrif	Gruppo Cer.che Ricchetti	Fmr - Art'e'
Pirelli & c. R.e.	Indesit	Gruppo editoriale l'espresso
Risanamento	Interpump	Il Sole 24 Ore
Ubi banca	Irce	Iride
Unipol	It holding	Mediaste
Vittoria	La Doria	Reply
Credito Valtellinese	Maffei	Sat
Kme Group	Mariella Burani Group	Snam Rete Gas
	Nice	Telecom
	Permasteelisa	
	Piaggio	
	Pininfarina	
	Poligrafica San Faustino	
	Poltrona Frau	
	Rdb	
	Ricordati	
	Sabaf	
	Saes Getters	
	Socotherm	
	Stefanel	
	Trevi	
	Vianini Ind	

Number of insulated elements in financial sector: 19: industrial: 33; services: 20

3.2 Assortativity

Assortativity is closely related to correlation of nodes degrees and it measures the tendency of high connected nodes to be connected with other high connected nodes. The more the assortativity is close to 1, the more high connected nodes are connected with other high connected nodes, hence networks show a big central cluster of high

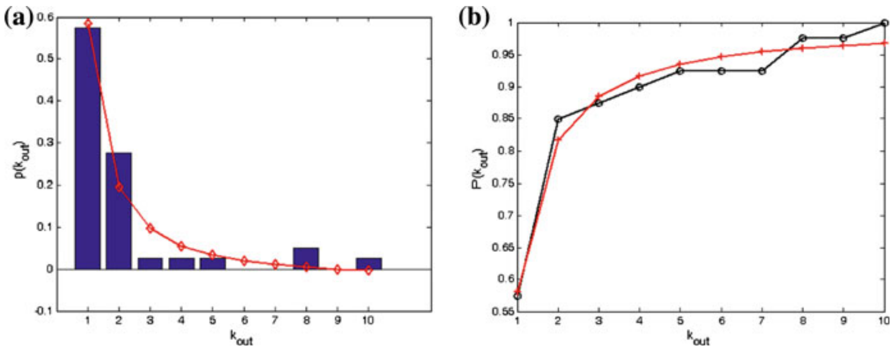

Fig. 5 Financial sector. **a** Histogram (*bars*) and power law best fit (continuous *grey curve* with diamonds). Estimate of $p(k_{\text{out}}) \sim k_{\text{out}}^{-\gamma}$ gives $\gamma = 1.467\,(0.5716, 2.368)$, RMSE$\,=\,$0.0477. The Jarque–Bera test validates the hypothesis of Gaussianity of residuals. **b** Empirical probability distribution (*black line* and *circles*) and power law best fit (continuous *grey curve* with diamonds). Estimate of $P(k_{\text{out}}) \sim k_{\text{out}}^{1-\gamma}$ gives $\gamma = 2.222\,(1.5912, 2.853)$, RMSE$\,=\,$0.0250. The MLE gives $\gamma = 2.64\,(2.35, 2.94)$, so leading to $\gamma\,(2.35)$. The Jarque–Bera test accepts the hypothesis of Gaussianity of residuals

connected nodes surrounded by satellites. The more the assortativity is close to -1, the more high connected nodes tend to be connected with low connected nodes, so the networks picture looks like a collection of star networks.

Therefore, assortativity measure contributes to analyze clusters in networks and to examine the connection among them.

Let through s_{ij} be the percentage of shares of firm j held by firm i. The adiacency matrix of our network is the n-square matrix

$$A = \left(s_{ij}\right)$$

representing the direct cross-ownership in the set. Either the presence or the absence of a link from i to j is given by $a_{ij} = \text{sign}(s_{ij})$, that is 1 if the link exists, and 0 if the link does not exist.

In our case, the assortativity has been calculated considering the direction of the links following Eq. (26) Newman (2003), that can be conveniently rewritten as

$$r = \frac{\sum_{i,j=1}^{N} q_i^{\text{in}} q_i^{\text{out}} a_{ij} - \frac{1}{L}\left(\sum_{i,j=1}^{N} q_i^{\text{in}} a_{ij}\right)\left(\sum_{i,j=1}^{N} q_i^{\text{out}} a_{ij}\right)}{\sqrt{\left[\sum_{i,j=1}^{N}\left(q_i^{\text{in}} a_{ij}\right)^2 - \frac{1}{L}\left(\sum_{i,j=1}^{N} q_i^{\text{in}} a_{ij}\right)^2\right]\left[\sum_{i,j=1}^{N}\left(q_i^{\text{out}} a_{ij}\right)^2 - \frac{1}{L}\left(\sum_{i,j=1}^{N} q_i^{\text{out}} a_{ij}\right)^2\right]}}$$

where, for each node i, k_{out}^i is the out-degree of node i, k_{in}^i is the in-degree of node i, p_i^{out} and p_i^{in} are the probabilities that a randomly chosen vertex i has out-degree (or in-degree, respectively) equal to k_{out}^i (k_{in}^i), and $q_i^{\text{out}} = \frac{k_{\text{out}}^i p_i^{\text{out}}}{\sum_i k_{\text{out}}^i p_i^{\text{out}}}$ and $q_i^{\text{in}} = \frac{k_{\text{in}}^i p_i^{\text{in}}}{\sum_i k_{\text{in}}^i p_i^{\text{in}}}$ are the excess out- and in-degree.

The assortativity on the entire dataset gives 0.865; this means that there is a strong tendency to form a high connected group.

Table 6 reports the assortativity coefficients inside and among sectors.

Table 6 Assortativity coefficients for financial, industrial and services macrosectors

Macrosectors	Financials	Industrials	Services
Financials	0.776	0.782	0.513
Industrials	0.732	0.998	0.740
Services	0.000	1.000	0.987

Let us enter into each of the group that we examined, the results of the analysis on the three macro-sectors reveals that all the assortativity coefficients are smaller than average when the shareholder is a financial company. Assortativity inside the group is 0.786, 0.782 and 0.513, respectively, with industrial and services. The output seems to confirm the trend of financial domestic companies to own smaller and more recently listed companies, as reported in Bianchi et al. (2008). More than 50% of Italian listed companies were listed after 1998; they are smaller with respect to "old" companies, and are controlled through a coalition rather than by a single agent.

The assortativity of firms belonging to the industrial sectors are 0.732, 0.998 and 0.740; these data reveals that the network originating from industrials has high coefficients, particularly if the subsidiary is an industrial company.

The services sector reveals the maximum positive assortativity (1.00 with industrials) and the minimum value of zero towards financial companies.[4]

3.3 Network hierarchy

The estimate of the probability distribution of k_{out} is only a first step for the network analysis. Issues on the resilience of networks to attacks, contagions, and other diffusive phenomena Newman et al. (2006) have shown the relevance of the aggregation of nodes. Therefore, we proceed estimating a hierarchy measure.

We follow the definition of hierarchy in accord with Trusina et al. (2004), where a node i has a level higher of node j if $k^i_{\mathrm{out}} > k^j_{\mathrm{out}}$, i.e. if the company i diversifies its portfolio more than the company j. A path between two nodes of a network is told "hierarchical" if it is composed by an "up" path (only links to more connected nodes are allowed) followed by a "down path" (only links to less connected nodes are allowed) (Trusina et al. 2004). Hierarchy measures the percentage of nodes that have a common ancestor, i.e. the percentage companies whose shares are owned by a third company with a more diversified portfolio. Of course, the length of the shortest hierarchical path between a given pair of nodes could not exist at all, if these two nodes cannot be connected by any hierarchical path. Otherwise, the length of the shortest hierarchical path between a given pair of nodes can be either equal to the length of the shortest path between these nodes or longer than it. The last case happens when

[4] The output of zero for the assortativity coefficient for the services companies owning financials is due to the fact that there are only 4 cases, from which we cannot infer a prevalent relationship: a) Autostrada To-MI ($k_{\mathrm{out}} = 1$) is connected to Assicurazioni Generali ($k_{\mathrm{in}} = 2$); b) Poligrafici Editoriale ($k_{\mathrm{out}} = 1$) is connected to Monrif ($k_{\mathrm{in}} = 1$); c) RCS Mediagroup ($k_{\mathrm{out}} = 1$) is connected to Intesa San Paolo ($k_{\mathrm{in}} = 1$); d) SIAS ($k_{\mathrm{out}} = 3$) is connected to Alerion ($k_{\mathrm{in}} = 1$), Assicurazioni Generali ($k_{\mathrm{in}} = 2$), Banca Carige ($k_{\mathrm{in}} = 1$).

there is non-hierarchical shortcut—a path shorter than the shortest hierarchical path between a pair of nodes.

Therefore, the number of hierarchical paths evidences to which extent a higher diversification corresponds to a higher potential control. The number of paths that are not hierarchical evidences shortcuts between nodes that do no respect the hierarchical structure. If the path is also a path of control, then the two companies can communicate mainly only through their common controller. In this section we report the analysis done considering the links only and not their weight, leaving for the next section the estimate of the amount of control paths that are also hierarchical.

On the data set it resulted that the length of 1549 hierarchical paths is equal to the length of the shortest path, and that 954 hierarchical paths have a length bigger than the length of the shortest path. The probability to find a shortcut not respecting the hierarchy is 0.62.

We performed a scenario analysis rewiring the network. Similar to a randomized version, the maximally hierarchical version of a network is generated by multiple rewiring pairs of edges. One has to add however a particular preference for reconnection: at each step one selects two pairs of connected nodes and attempts to connect the node with the highest among these four nodes to the node of the next highest degree in this subset. The remaining two nodes are then linked together. Multiple links are forbidden.

The maximally anti-hierarchal version of a network can be constructed by the same algorithm but with the opposite preference of reconnection: the node with the highest degree is linked with that with the lowest degree.

On 100 maximally hierarchical randomized networks the percentage was in the range (0.4, 0.9). On 100 minimally hierarchical randomized networks the percentage was in the range (0.0, 0.7). Therefore, the network is close to the mean in the case of maximally hierarchical networks, but it has a high level of hierarchy whether compared to minimally hierarchical networks. Therefore, we can conclude that the hierarchy in the network is quite relevant.

3.4 The relationship between diversification and dimension

The existence of a link from i to j shows that company i holds shares of company j. We want to get deeper into the analysis of diversification, studying the relation between the degree of diversification and several dimensional parameters describing the size of both the controller and the controlled company.

In order to perform this analysis, we assign to each link from node i to node j the weight s_{ij} given by the percentage of shares of j[5] held by i. The dimensional parameters are the following:

– $c = (c_i)_{i \in E}$ is market capitalization[6] of the companies assigned to nodes. The product of the percentage s_{ij} and the market capitalization c_j of asset j give the

[5] Number of shares j held/total number of outstanding shares of asset j.

[6] The data of market capitalization of listed companies were obtained through Thomson Datastream database.

Table 7 Correlation coefficients between the number of outgoing arches k, the market value of subsidiaries v, the shareholder capitalization c and the holder total asset ta, and related quantities

	k v	c	ta	v^c	v^{ta}
a.1 Pearson					
k	0.46 (0.36, 0.55)	0.42 (0.31, 0.52)	0.45 (0.34, 0.54)	0.23 (0.11, 0.35)	0.11 (−0.2, 0.23)
v		0.69 (0.62, 0.75)	0.25 (0.13, 0.37)	0.30 (0.18, 0.41)	0.20 (0.08, 0.32)
c			0.72 (0.65, 0.77)	0.00 (−0.13, 0.12)	0.02 (−0.11, 0.14)
ta				0.01 (−0.12, 0.13)	−0.02 (−0.14, 0.11)
v^c					0.71 (0.65, 0.77)
v^{ta}					
a.2 Spearman					
k	0.92 (0.90, 0.94)	0.35 (0.24, 0.46)	0.38 (0.27, 0.48)	0.90 (0.88, 0.93)	0.89 (0.86, 0.91)
v		0.45 (0.35, 0.55)	0.46 (0.35, 0.55)	0.98 (0.97, 0.98)	0.96 (0.95, 0.97)
c			0.85 (0.81, 0.88)	0.35 (0.24, 0.46)	0.33 (0.21, 0.43)
ta				0.37 (0.26, 0.47)	0.31 (0.19, 0.42)
v^c					0.98 (0.98, 0.99)
v^{ta}					
a.3 Kendall					
k	0.56 (0.0001)	0.28 (0.0001)	0.30 (0.0001)	0.54 (0.0001)	0.52 (0.0001)
v		0.36 (0.0001)	0.36 (0.0001)	0.57 (0.0001)	0.54 (0.0001)
c			0.68 (0.0001)	0.26 (0.0001)	0.24 (0.0001)
ta				0.27 (0.0001)	0.22 (0.0001)
v^c					0.59 (0.0001)
v^{ta}					

The statistics refer to entire data sets. Parentheses indicate the 95% confidence interval (Pearson and Spearman), and the 95% threshold for the acceptance for the existence of non-zero correlation (Kendall)

market value of the investment of i in j, the market capitalization of the companies;

- $v = (v_i)_{i \in E}$. v, where $v_i = \sum_{j \in J} s_{ij} c_j$ and J is the set of nodes neighboring i, gives the market value of the entire portfolio of subsidiaries of i, i.e. the total amount of wealth of i invested in other companies;
- ta is the total asset indicator of Datastream; it is used as a *proxy* for firm size, and it measures the capital employed in the activity, the sum of short, medium and long term investment asset;
- $v^c = (v_i^c)_{i \in E}$, where $v_i^c = v_i / c_i$ is a ratio that displays the weight of the investment in other listed companies relative to the dimension of the holder, measured through capitalization;
- $v^{ta} = (v_i^{ta})_{i \in E}$, where $v_i^{ta} = v_i / c_i$ is a ratio that displays the weight of the investment in other listed companies relative to the dimension of the holder, measured through total asset.

We calculate both rank order correlation coefficient (Spearman and Kendall) and Pearson correlation coefficient, for the entire sample (Table 7) and for the three sectorial subsamples (Tables 8 for financials, Table 9 for industrials, and Table 10 for

Table 8 Correlation coefficients between the number of outgoing arches k, the market value of subsidiaries v, the shareholder capitalization c and the holder total asset ta, and related quantities

	k v	c	ta	v^c	v^{ta}
b.1 Pearson					
k	0.62 (0.45, 0.75)	0.54 (0.34, 0.69)	0.38 (0.15, 0.57)	0.11 (−0.13, 0.35)	0.03 (−0.22, 0.27)
v		0.37 (0.14, 0.56)	0.25 (0.00, 0.46)	0.41 (0.19, 0.60)	0.22 (−0.02, 0.44)
c			0.96 (0.94, 0.98)	−0.08 (−0.32, 0.16)	−0.05 (−0.29, 0.19)
ta				−0.09 (−0.32, 0.16)	−0.07 (−0.30, 0.18)
v^c					0.69 (0.54, 0.80)
v^{ta}					
b.2 Spearman					
k	0.92 (0.87, 0.95)	0.52 (0.31, 0.68)	0.37 (0.14, 0.56)	0.84 (0.75, 0.90)	0.81 (0.70, 0.88)
v		0.57 (0.38, 0.71)	0.44 (0.23, 0.62)	0.92 (0.87, 0.95)	0.87 (0.79, 0.92)
c			0.79 (0.68, 0.87)	0.32 (0.09, 0.52)	0.24 (0.00, 0.46)
ta				0.20 (−0.05, 0.42)	0.06 (−0.19, 0.30)
v^c					0.96 (0.94, 0.98)
v^{ta}					
b.3 Kendall					
k	0.70 (0.0001)	0.40 (0.0001)	0.27 (0.0001)	0.59 (0.0001)	0.57 (0.0001)
v		0.43 (0.0001)	0.32 (0.0001)	0.69 (0.0001)	0.61 (0.0001)
c			0.65 (0.0001)	0.23 (0.0089)	0.16 (0.0628)
ta				0.14 (0.1190)	0.04 (0.6807)
v^c					0.76 (0.0001)
v^{ta}					

The statistics refer to financial sub-sample. Parentheses indicate the 95% confidence interval (Pearson and Spearman), and the 95% threshold for the acceptance for the existence of non-zero correlation (Kendall)

services). Parentheses indicate the 95% confidence interval (Pearson and Spearman), and the 95% threshold (Kendall) for the acceptance for the existence of non-zero correlation. In few cases, a relationship found significant with rank correlation coefficients is not supported by the Pearson correlation coefficient. In these cases, we prefer to consider the non parametric indicators for the less severe assumptions required.

We first tried to determine whether the companies that most diversify the portfolio (having the highest k) also have the biggest amount of capital invested in other firms. The correlation between k and v is 0.46 for the whole sample. This outcome proves that the level of diversification is highly correlated to the amount of wealth invested in other companies. Therefore, there is a good agreement between portfolio diversification and wealth, i.e., actually, the companies that most diversify their portfolio are also the ones with highest wealth invested. The results confirm previous findings, showing a peak of the correlation coefficient for the financial branch and a minimum for the industrial subsample.

Moreover, we tried to link the degree of diversification to the size of the holder. The first test aims to understand whether high cap companies tend to diversify more

Table 9 Correlation coefficients between the number of outgoing arches k, the market value of subsidiaries v, the shareholder capitalization c and the holder total asset ta, and related quantities

k	v	c	ta	v^c	v^{ta}
c.1 Pearson					
k	0.36 (0.19, 0.52)	0.36 (0.19, 0.51)	0.44 (0.27, 0.58)	0.55 (0.40, 0.67)	0.64 (0.52, 0.74)
v		0.98 (0.96, 0.98)	0.86 (0.80, 0.90)	0.22 (0.03, 0.39)	0.52 (0.37, 0.65)
c			0.93 (0.90, 0.95)	0.18 (−0.01, 0.36)	0.48 (0.32, 0.62)
ta				0.20 (0.01, 0.37)	0.44 (0.27, 0.58)
v^c					0.80 (0.73, 0.86)
v^{ta}					
c.2 Spearman					
k	0.89 (0.84, 0.92)	0.18 (−0.01, 0.36)	0.22 (0.03, 0.39)	0.88 (0.83, 0.92)	0.88 (0.83, 0.92)
v		0.31 (0.13, 0.47)	0.28 (0.10, 0.45)	0.98 (0.98, 0.99)	0.98 (0.97, 0.99)
c			0.85 (0.79, 0.90)	0.24 (0.05, 0.41)	0.24 (0.06, 0.41)
ta				0.23 (0.04, 0.40)	0.20 (0.01, 0.37)
v^c					0.99 (0.99, 0.99)
v^{ta}					
c.3 Kendall					
k	0.45 (.0001)	0.22 (0.0195)	0.22 (0.0167)	0.44 (0.0001)	0.44 (0.0001)
v		0.29 (0.0012)	0.28 (0.0020)	0.48 (0.0001)	0.47 (0.0001)
c			0.78 (0.0001)	0.21 (0.0190)	0.21 (0.0221)
ta				0.20 (0.0273)	0.19 (0.0355)
v^c					0.51 (0.0001)
v^{ta}					

The statistics refer to industrial sub-sample. Parentheses indicate the 95% confidence interval (Pearson and Spearman), and the 95% threshold for the acceptance for the existence of non-zero correlation (Kendall)

(both in number of different assets and wealth) than small cap ones. The analysis of the correlation between k and the shareholder capitalization c shows similar results (positive correlation with a peak on financial companies), proving a stronger degree of diversification for high-cap holding companies.

For the industrials stocks zero is inside the confidence interval of the Spearman correlation among k and c, while the Pearson correlation is positive; this evidence shows that the values of the capitalization and diversification are much closer than the corresponding ranks, so showing an independence among the policies of diversification and the size in terms of capitalization of the company. This means that most industrial companies follow their diversification strategy nearly independently from their size. Analogously, but in the opposite manner, for the services sector the correlation between k and c is significant only when calculated on ranks; again, the companies' size, in terms of capitalization, is not relevant to their portfolio diversification, i.e. ownership strategies on the market are not only due to the possibility to buy shares of other companies. The analysis is very similar when k is linked to total asset ta: the non significant Pearson coefficient improves when computed on ranks.

Table 10 Correlation coefficients between the number of outgoing arches k, the market value of subsidiaries v, the shareholder capitalization c and the holder total asset ta, and related quantities

k	v	c	ta	v^c	v^{ta}
d.1 Pearson					
k	0.41 (0.20, 0.59)	0.21 (−0.03, 0.42)	0.18 (−0.05, 0.39)	0.25 (0.02, 0.45)	0.28 (0.05, 0.48)
v		0.30 (0.07, 0.49)	0.29 (0.06, 0.49)	0.77 (0.66, 0.85)	0.69 (0.54, 0.79)
c			0.97 (0.95, 0.98)	−0.03 (−0.26, 0.20)	−0.04 (−0.26, 0.20)
ta				−0.02 (−0.25, 0.21)	−0.03 (−0.26, 0.20)
v^c					0.92 (0.88, 0.95)
v^{ta}					
d.2 Spearman					
k	0.88 (0.82, 0.92)	0.28 (0.05, 0.48)	0.28 (0.06, 0.48)	0.87 (0.81, 0.92)	0.87 (0.80, 0.92)
v		0.38 (0.16, 0.56)	0.36 (0.14, 0.55)	0.98 (0.96, 0.99)	0.97 (0.96, 0.98)
c			0.92 (0.87, 0.95)	0.28 (0.06, 0.48)	0.28 (0.06, 0.48)
ta				0.27 (0.04, 0.47)	0.25 (0.02, 0.46)
v^c					1 (0.99, 1)
v^{ta}					
d.3 Kendall					
k	0.45 (0.0001)	0.22 (0.0195)	0.22 (0.0167)	0.44 (0.0001)	0.44 (0.0001)
v		0.29 (0.0012)	0.28 (0.0020)	0.48 (0.0001)	0.47 (0.0001)
c			0.78 (0.0001)	0.21 (0.0190)	0.21 (0.0221)
ta				0.20 (0.0273)	0.19 (0.0355)
v^c					0.51 (0.0001)
v^{ta}					

The statistics refer to services sub-sample. Parentheses indicate the 95% confidence interval (Pearson and Spearman), and the 95% threshold for the acceptance for the existence of non-zero correlation (Kendall)

The analysis of macrosectors reveals that the low values of Pearson correlation coefficient between k and both v^c and v^{ta} (the last estimate is non significant) greatly improve for the industrials: for the sector the link of k and v is weak, but it increases when v is divided by a dimension factor. It is interesting to note that all these estimates are positive and strongly significant when measured on ranks.

The high value of the (Pearson) correlation coefficient between v, i.e. the wealth invested in other companies participating their capital, and c, the capitalization of the holder suggests a strongerconnection between the size of the firm and the wealth invested: high cap companies diversify more than small cap.

The result for the whole sample is raised by the almost perfect coefficient of the industrial sub-sample: for this sector, the proposition "bigger dimension, bigger investment in other companies" is a perfect rule, and it depends only in part on the number of arches k. The outcome might be due to the attitude of industrial firms to concentrate in the mere entrepreneurial activity; in this context the goal of diversification, pursued through the acquisition of shareholdings, is achieved only in presence of specific dimensional conditions. Nevertheless the coefficients of the other areas are significant, and the results are confirmed, on a weaker basis, with respect to ta instead of c.

For the remaining coefficients of the entire sample, not only k but also c is not correlated with v^{ta}, and the absence of correlation is confirmed by the correlation among ta and v^c. The comparison of this result with the positive correlation among k and v, c and v, k and ta, c and ta shows that ta gets all the relevant information for c. The zero correlation among $c$and v^c, together with the positive correlation among c and v shows that c gets the biggest part of the information contained in v. The low values of the correlation among v and v^c, and among v and v^{ta} show that c and ta play a relevant role for the information contained in$v$. This is straightforward due to the formula connecting v and c, but it is not trivial when comparing v and ta. This means that ta plays a role quite close to c. Finally, for the whole sample the cases of no significance for Pearson correlation coefficient are totally reverted when considering the rank correlation coefficient (both Spearman and Kendall): this means that ranks are not changing, although the values are quite different.

In the financial sector, for Pearson correlation coefficients between k and v^c, and v and v^{ta} the hypothesis of zero correlation cannot be rejected. This means that the diversification of the portfolio, represented by k, in the build. up of v^c in the financial sector is much more relevant than on the entire data set. This confirms the higher values of k in the financial sector, confirming the main role of some financial companies as managers of the financial part of other companies. This fact is as relevant as causing the lower values in the Spearman ranks, till to allowing zero Spearman rank correlation among ta and v^c, and among ta and v^{ta}. This means that the capitalization is highly due to the total asset, and that it captures most of the information from v, so behaving quite like the capitalization. The increase of the significance threshold in the Kendall rank correlation confirms the above results.

For the industrials, the relevance of the link between v and c produces the insignificance of the Pearson correlation coefficient among c and v/c, proving that the capitalization in our industrial sample is the true proxy of the diversification.

Finally, in the services sector the capitalization has a big impact on v, meaning that companies in the services sector are not keen to buy shares of other companies of the same sector, and this can be deduced from the (Pearson) correlation among c and v^c. The Pearson correlation coefficient among $c$and v^{ta}, and among ta and v^c, together with the high values of the correlation coefficient among c and ta shows the closeness of c and ta.

Spearman rank correlation and Kendall rank correlation coefficient confirm the results obtained through the Pearson correlation coefficient.

4 Control and integrated ownership

In this section we follow the approach outlined in Chapelle and Szafarz (2005) looking for the paths of control and for the integrated ownership, i.e. for the ownership that is obtained through intermediary companies. The presence of cycles in our network prevents the usage of the algorithm shown in Salvemini et al. (1995), nevertheless conditions on the norm of the adjacency matrix ensure that all the relevant information is caught.

4.1 Paths of control

We aim to estimate the paths composed by edges whose weight is bigger than 50%, to emphasize the direct control of a company onto another one.

We also want to detect chains of control. We assume that company A controls company B if there is a chain of ownership, corresponding to the existence of a path on the network starting from A and ending in B such that the weight on each link is more than 50%. This means that A has the majority on the Board of Directors, and that it can drive the decisions of B that request the simple majority of votes. A similar analysis can be done for the 66% vote's level.

The control of A on B can be achieved also if A has a very low level of shares of B. In fact, in case of a chain of ownership, the final percentage of shares of B held by A are given by the product of the intermediate ownerships, that go to zero quite fast because at each step there is the product of numbers lower than 1. As example, if A owns 51% of C, that owns 51% of D, that owns 51% of E, that owns 51% of F, that owns 51% of B, it results that A controls B, and A owns only the $51\%(51\%(51\%(51\%(51\%)))) = 3{,}45\%$ of B. A company G owning directly the 5% of B has a higher ownership, but no control.

This example shows that the control of a firm is very different from the ownership of it. It also could allow creating mechanisms to control many firms by using a very limited amount of funds. Our sample has a very few paths of controls, that are shown in Table 11. There are 16 companies that control directly other 16 companies. Controllers are most financial companies. There are other three financial companies that control 2 companies each. The usual length of the chains of control is 1. It is 2 only for Ifi Priv-Ifil-Juventus.

We are interested in understanding whether there is some relationship among the number of controlled companies and out-degree. Therefore we want to know whether the companies that most diversify their portfolio also achieve the control of some other company. The correlation coefficient of the number of controlled companies and out-degree shows a mild dependence. We also get that in 21 of the 24 cases reported, companies with bigger out-degree control companies with lower out-degree. This is most due to the presence of financial companies, that are the ones with the highest out-degree.

We also consider the correlation of the number of controlled companies and v. The result shows that diversification and wealth invested increase nearly in the same way.

We repeat the analysis considering the capitalization. The correlation coefficient of the number of controlled companies shows that the pursuit of control is not only a property of high cap companies.

The analysis of direct control must be compared with integrated ownership and integrated control, that are the matter of the next sections.

4.2 Integrated ownership

Ownership data can be further examined to detect complex chains of ownership got through intermediate companies, i.e. indirect ownership. In fact, due to

Table 11 Links whose weight is higher than 0.5

	Controller	Controlled	% of shares
Section a)			
Financials	Ass.ni Generali	Bca Generali	0.598
	Banca Popolare	Cred Bergamasco	0.878
	Bastogi	Brioschi	0.533
	Immsi	Piaggio	0.559
	Intesa San Paolo	Sirti	0.940
	Italmobiliare	Italcementi	0.603
	Monrif	Poligrafici Ed	0.600
	Pirelli & C.	Pirelli & c. R.E.	0.550
	Credito Valtell.se	Credito Artigiano	0.655
	Intek	Kme Group	0.538
Industrials	Cremonini	Marr	0.574
	Eni	Snam Rete Gas	0.500
	Erg	Enertad	0.684
Services	Autostr. TO-MI	Sias	0.617
	Iride	Mediterranea	0.683
	Telecom	Telecom It.Media	0.669
Section b)			
	Caltagirone	Vianini	0.541
	Caltagirone	Vianini Lavori	0.500
	Cir	Gr. Ed. Espresso	0.509
	Cir	Sogefi	0.583
	Fondiaria - Sai	Immob. Lombarda	0.789
	Fondiaria - Sai	Milano	0.511
Section c)			
	Ifi Priv	Ifil	0.668
	Ifil	Juventus	0.620

(a) companies that control just another company. (b) companies that control 2 other companies. In this case the controllers are all financial companies. (c) The only chain: Ifi Priv controls Ifil, that controls Juventus

cross-shareholding, however, the true ownership structure may be hidden by a complex network of indirect relations.

In this part of the paper we most base on Chapelle and Szafarz (2005), that starts form an input-output matrices approach, and develop a method that evidence the final owner. A similar analysis is carried on control relationships. Techniques from operational research are used, considering again the paths on the network, and emphasizing the difference between ownership and control. In order to explain how indirect control may rise let us consider the situation outlined in Fig. 6. Company A is assumed to control company B whether A owns more than 50% of shares of B, either directly (*direct ownership*) or through intermediate companies (*integrated ownership)*. Company A owns directly 20% of shares of B. But A has also 60% of D, that has 30% of B, therefore A has 30% (60%) of B through D. Analogously, A has 30% of C that has 40% of B, therefore A has another 40%(30%) of B through C. Last but

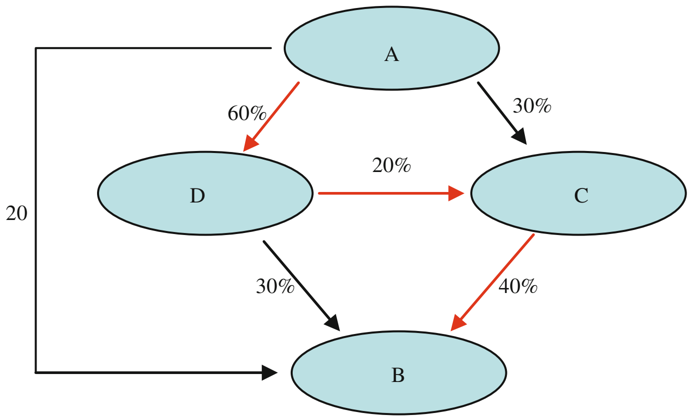

Fig. 6 Estimate of ownership

not least, there is another path for A to arrive to B: through D and then C, corresponding to another 40%(20%(60%)). The integrated ownership of A in B is the sum 20% + 40%(30%) + 30%(60%) + 40%(20%(60%)) = 54.8%. Therefore, A actually owns more than 50% of B, although the direct ownership is 20%.

The integrated ownership can be calculated by summating the weights of the paths that lead from A to B, where the weight of the path is given by the multiplication of the weights of the links in the path. The algorithm proposed in Salvemini et al. (1995) works under the hypothesis that the shareholding network is a directed acyclic graph. This hypothesis is not valid in our case.

A first remark consists in the fact that a path analysis is possible due to the fact that edge weights are in (0,1). In fact, the weight of longer paths goes quickly to zero, because anytime an edge more is considered, the entire path is multiplied by a number lower than one. Therefore, cycles visiting a node give a very low contribution to the integrated ownership. Remembering that the weights of paths of length n are given by the n-power of the adjacency matrix, we have that its norm goes very quickly to zero. In our sample the norm of the adjacency matrix is lower than 10^{-4} if $n > 7$.

Other papers Chapelle and Szafarz (2005), Dietzenbacher and Temurshoev (2008), Flath (1992), Huber and Ryll (1989), Turnovec (1999), Turnovec (2005) propose an input-output approach for the analysis that overcomes the problem of the presence of cycles, including the above remark as a particular case. The approach most suitable for our analysis is one developed in Chapelle and Szafarz (2005). In fact, while Huber and Ryll (1989) is the precursory paper of all that series, Turnovec (1999, 2005) use the hypothesis that the ownership sum up to 1; moreover Turnovec (1999, 2005) make the difference between primary and secondary owners. Primary owners correspond to sources in the network, and they are identified with individuals or Institutions. While "primary owners" can be matched to "final owners" in Chapelle and Szafarz (2005), this matching cannot go further on our dataset, because it is made of companies traded in the Stock Market. Moreover, the analysis in Chapelle and Szafarz (2005) evidences the role of secondary owners in the network, since primary owners do not manage completely secondary owners. Remembering that we indicated through s_{ij} the percentage of shares of firm j held by firm i, we have that the n-square matrix

$$A = \left(s_{ij}\right)$$

represents the direct cross-ownership in the set. The fact that secondary owners maintain some freedom can be resumed into the condition

$$\sum_{j=1}^{n} s_{ij} \leq 1, \quad i = 1, \ldots, n$$

Integrated ownership is the sum of all direct and indirect participations. The sum is performed on all the paths of the network. Remembering that the power k of matrix A represents the weight of the paths of length k Chapelle and Szafarz (2005), Turnovec (1999, 2005), we have that

$$Y = \sum_{j=1}^{\infty} A^k = (I - A)^{-1} A$$

contains the weights among each possible couple of nodes The main hypothesis of the existence of the inverse of the matrix (I-A) holds in our case. The computation of integrated ownership is not obvious because the simple sum of direct and indirect participations suffers from a double counting problem. Indeed, when adding up all levels of ownership, the same shares are considered at each ownership level. Therefore, the sum y_{ij} of all shares held by direct and indirect shareholders of a company often exceeds 100%. The integrated ownership matrix can be cleaned from double counting, giving rise to the following matrix

$$V = (diag(I - \overline{A}))Y, \quad \text{where} \quad \overline{A} = \left(\sum_{i=1}^{n} a_{ij}\right)$$

In order to clarify the method let us show another example and consider matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0.4 \\ 0 & 0 & 0 & 0.6 & 0 \\ 0 & 0 & 0 & 0.4 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nodes 2 and 3 own the 100% of node 4, that owns 50% of node 5. Therefore, matrix V allocates to the primary owners all the ownership of unit 4:

$$V = \begin{pmatrix} 0 & 0 & 0 & 0 & 0.4 \\ 0 & 0 & 0 & 0.6 & 0.3 \\ 0 & 0 & 0 & 0.4 & 0.2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

On this example, matrix Y differs from V only for the presence of the link from node 4 to node 5, weighting 0.5. Figure 7 shows matrices A and V.

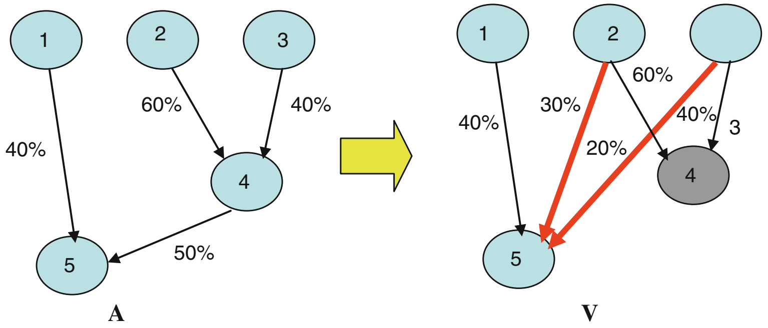

Fig. 7 Integrated ownership: the case of a secondary owner completely owned

If node 4 is not completely owned by nodes 2 and 3, as example

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0.4 \\ 0 & 0 & 0 & 0.6 & 0 \\ 0 & 0 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

then a percentage is left and

$$V = \begin{pmatrix} 0 & 0 & 0 & 0 & 0.4 \\ 0 & 0 & 0 & 0.6 & 0.3 \\ 0 & 0 & 0 & 0.4 & 0.15 \\ 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Figure 8 shows matrices A and V. Table 12 reports indirect ownership that allows firm i to own more than 0.5 of firm j. We note that in most cases the integrated ownerships overlaps with control, although the percentage of share may change. There are a few companies that go below the 0.5 threshold, due to the reallocation of ownership to the

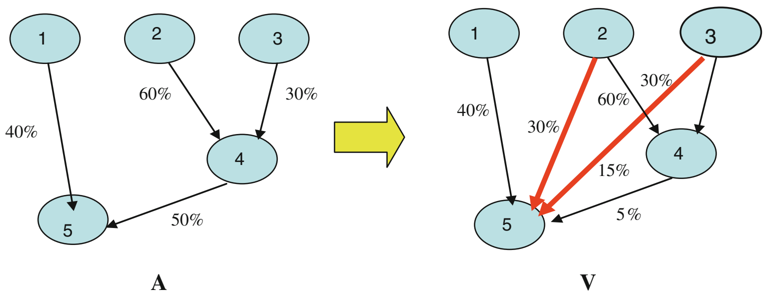

Fig. 8 Integrated ownership: the case of a secondary owner not completely owned

Table 12 Comparison between Table 11 and integrated ownership

	Controller	Controlled	% Direct control	% Integr. ownership	
Section a)					
Financials	Ass. Generali	B.ca Generali	0.598	<0.5	
	B. Popolare	Cr Bergamasco	0.878	0.878	Identical
	Bastogi	Brioschi	0.533	0.533	Identical
	Immsi	Piaggio	0.559	0.569	Higher
	Intesa S.P.	Sirti	0.940	0.881	Lower
	Italmobiliare	Italcementi	0.603	0.564	Lower
	Monrif	Poligrafici Ed	0.600	0.602	Higher
	Pirelli & C.	Pirelli & C.R.E.	0.550	<0.5	
	Cr. Valtell.se	Cr.to Artigiano	0.655	0.655	Identical
	Intek	Kme Group	0.538	0.538	Identical
	Intek	Greenenergy C.	<0.5	0.518	New
Industrials	Cremonini	Marr	0.574	0.574	Identical
	Eni	Snam Rete Gas	0.500	0.541	Higher
	Erg	Enertad	0.684	0.684	Identical
Services	Auto TO-MI	Sias	0.617	0.58	Lower
	Iride	Mediterranea	0.683	0.683	Identical
	Telecom	Telecom It. M.	0.669	0.647	Lower
Section b)					
	Caltagirone	Vianini	0.541	0.541	Identical
	Caltagirone	Vianini Lavori	0.500	0.500	Identical
	Cir	Gr.Ed. Espresso	0.509	<0.5	
	Cir	Sogefi	0.583	<0.5	
	Fond. Sai	Imm.Lombarda	0.789	<0.5	
	Fond. Sai	Milano	0.511	<0.5	
Section b)					
	Ifi Priv	Ifil	0.668	0.676	Higher
	Ifil	Juventus	0.620	<0.5	

final owner, and a new entry in the table only, due to the relationship between Intek and Greenenergy capital. We also get that in 15 of the 18 cases reported, companies with bigger out-degree control companies with lower out-degree.

4.3 Effective control

In the previous section we examined integrated ownership, direct control, and control achieved through chains of controlled companies.

Now we are interested in knowing the relative majority, and in understanding whether it is due either to complex patterns or most to direct control.

Effective control is obtained by applying the majorization rule to the matrix of voting rights, that we identify with matrix A. A voting share bigger than the half of shares is turning into absolute control (100%), meaning that firm I holds full control over firm j. Moreover, since the other shareholders of firm j are expropriated from control, their effective control over firm j is set equal to zero.

We also estimate the owner that has the highest percentage of integrated ownership in each company, not necessarily bigger than 50%, and we estimate the correlation coefficient between such percentage and the out-degree. In fact, a company might not have the strict majority, but a good level of shares may be sufficient for fast agreement with other counterparts in order to reach the majority.

Therefore, it is interesting to know whether the companies that most diversify their portfolio maintain the relative majority, so that they have some relevance in control and alliances. This definition of effective control formalizes the expropriation faced by the minority shareholders, largely documented in the literature, and it is estimated by the square matrix

$$c_{ij} = \begin{cases} 1 & \text{if } s_{ij} > 0.5 \\ 0 & \text{if } \exists k \neq i : s_{kj} > 0.5 \\ s_{ij} & \text{otherwise} \end{cases}$$

The transition from effective control to integrated control is similar to the transition from direct ownership to integrated ownership:

$$D = (diag(I - \overline{C}))(I - C)^{-1}C, \quad \text{where} \quad \overline{C} = \left(\sum_{i=1}^{n} c_{ij}\right)$$

We get that in the Italian Stock Market the chains of control are quite short, and that they most coincide with the direct ownership. This means that in our sample portfolio diversification is not an instrument that is used for dominating the market. Therefore, the situation is quite different from the *Keiretsu* Japanese structure of cross-ownership among companies. Moreover, there are financial companies that own and control just another company. Typically, they were created to be the financial part of a main company, and such kind of ownership is quite obvious. Table 13 shows and compare tables of ownership and control considering the two different majorization rules.

We also get that in 22 of the 25 cases reported, companies with bigger out-degree control companies with lower out-degree.

It results that in our sample the control is reached through direct ownership, and not through hidden intermediate Chinese boxes companies. In most cases, a financial society treats the financial part of another company. Paths are very short: one edge only in most cases, and a 2 edge path in the case of the financial companies of the football company "Juventus".

We also note that only 3 (B. Popolare, Bastogi, Caltagirone) out of the 16 sources (Table 3) actually control other companies.

Table 13 Comparison between Table 12 and integrated control

Controller	Controlled	% Direct control	% Integrated ownership	% Integrated control
Section a)				
Financials				
Ass.ni Generali	Bca Generali	0.598	<0.5	0.808
Banca Popolare	Cr. Bergamasco	0.878	0.878	1
Bastogi	Brioschi	0.533	0.533	1
Immsi	Piaggio	0.559	0.569	1
Intesa SP	Sirti	0.940	0.881	0.937
Italmobiliare	Italcementi	0.603	0.564	0.9
Monrif	Poligrafici Ed	0.600	0.602	1
Pirelli & C.	Pirelli & C. R.E.	0.550	<0.5	0.663
Cr.Valtellinese	Cred. Artigiano	0.655	0.655	1
Intek	Kme Group	0.538	0.538	1
Intek	Greenenergy C.	<0.5	0.518	0.550
Industrials				
Cremonini	Marr	0.574	0.574	1
Eni	Snam Rete Gas	0.500	0.541	1
Erg	Enertad	0.684	0.684	1
Services				
Autostr. TO-MI	Sias	0.617	0.58	0.952
Iride	Mediterranea	0.683	0.683	1
Telecom	Telecom It. M.	0.669	0.647	0.967
Section b)				
Caltagirone	Vianini	0.541	0.541	1
Caltagirone	Vianini Lav	0.500	0.500	1
Cir	Gr. Espresso	0.509	<0.5	0.522
Cir	Sogefi	0.583	<0.5	0.522
Fondiaria Sai	Imm.Lombarda	0.789	<0.5	0.611
Fondiaria Sai	Milano	0.511	<0.5	0.611
Section c)				
Ifi Priv	Ifil	0.668	0.676	1
Ifil	Juventus	0.620	<0.5	1

5 The giant component

We add a final remark on the giant component, since it is the one that contains cycles, and the analysis on the smaller component is straightforward. We report the correlations among capitalization, out-degree, wealth, ownership, control.

Results are shown in Table 14. While it is obvious that wealth and capitalization are highly correlated, we also get that capitalization and out-degree are positively cor-

Table 14 Correlation coefficients among capitalization, out-degree, wealth, ownership, control

(a) Pearson			Diversification		Network structure	
		Capitalization	k	Wealth	Ownership	Control
	Capitalization		***0.46*** *(0.29,0.60)*	***0.33*** *(0.15,0.50)*	***0.14*** *(-0.05,0.03)*	***0.15*** *(-0.04,0.34)*
Diversification	k			***0.64*** *(0.50,0.74)*	***0.16*** *(-0.03,0.35)*	***0.39*** *(0.21,0.54)*
	Wealth				***0.20*** *(0.01,0.38)*	***0.37*** *(0.18,0.52)*
Network structure	Ownership					***0.73*** *(0.63,0.81)*
	Control					

(b) Spearman			Diversification		Network structure	
		Capitalization	k	Wealth	Ownership	Control
	Capitalization		***0.48*** *(0.31,0.62)*	***0.54*** *(0.34,0.67)*	***0.14*** *(-0.06,0.33)*	***0.23*** *(0.04,0.41)*
Diversification	k			***0.92*** *(0.88,0.94)*	***0.28*** *(0.09,0.45)*	***0.42*** *(0.25,0.57)*
	Wealth				***0.40*** *(0.22,0.55)*	***0.51*** *(0.35,0.64)*
Network structure	Ownership					***0.81*** *(0.74,0.87)*
	Control					

(c) Kendall			Diversification		Network structure	
		Capitalization	k	Wealth	Ownership	Control
	Capitalization		***0.37*** *(0.0001)*	***0.41*** *(0.0001)*	***0.11*** *(0.1639)*	***0.19*** *(0.0212)*
Diversification	k			***0.65*** *(0.0001)*	***0.15*** *(0.0046)*	***0.24*** *(0.0001)*
	Wealth				***0.22*** *(0.0001)*	***0.31*** *(0.0001)*
Network structure	Ownership					***0.81*** *(0.0001)*
	Control					

(a) Pearson correlation coefficients (b) Spearman rank correlation coefficients (c) Kendall rank correlation coefficients

related. This means that the wealthy companies are the most keen to diversify their portfolios. From the positive correlation among out-degree and wealth we also get that the companies that most diversify their portfolio are also the ones with highest wealth invested. From the positive correlation among capitalization and ownership we also get that also capitalization follows the number of other firms owned, and the same remark holds for control. Continuing the analysis, the positive correlation among the out-degree and ownership reveals that many companies with a few links own other ones. The high correlation between out-degree and control is explained by the fact that some companies have the only role to be the financial company of a group. The positive, but low, correlation among wealth and ownership shows that out-degree is relevant for ownership much more than the total wealth, the same remark holds for control, too.

Such results are in line with the results already got on the entire data set. Spearman and Kendall rank correlation confirm the results obtained through the Pearson correlation.

6 Conclusions

The present report uses an interdisciplinary approach for the study of ownership and control in the Italian stock market. Both complex systems approach and operations research approach contribute to point out different elements of the shareholding network. The analysis shows that the shareholding network outlined in Sect. 2 reveals a structure having a giant weakly connected component with 101 nodes and 187 links and 7 smaller size weakly connected components with 2 or 3 nodes only. Considering the small sample size, there is a relatively high number of 0-connected companies (isolated nodes) and 1-connected companies. Focusing on the most connected companies, we show that the majority of them belongs to the financial sector. This result confirms the findings of Bianchi et al. (2008) that points out an increasing role of institutional investors in domestic companies' ownership in the last 15 years. The conclusion is that the diversification—which has to be considered in quantitative terms- of the portfolio of owned companies is attributable mostly to practical managerial issues rather than being used to gain control. Such a simple pattern emerged through quantitative analyses and measures breaks away from tunnelling outlined in the Japanese market, and it opens up new prospects in the literature on ownership structure, that documents an intensive use of pyramidal groups and very high ownership concentration and cross-holding in Continental Europe countries and in Italy (Aganin and Volpin 2005; Bianchi et al. 2008; Volpin 2002). Our study distinguishes from corporate governance works cited above for the methodological approach that abstracted from the particular case study of each group of companies, showing the results that can be achieved using public available information on the whole market rather than using specific economic and managerial knowledge on pyramidal structures, ultimately controlled by family groups. Analogously, our work distinguishes from the approach used in studies from the physics side like as Chapelle and Szafarz (2005), Garlaschelli et al. (2005) because we added knowledge to the mere analyses of the network structure examining the economic/financial relevance of companies in terms of integrated ownership and control. A deeper insight in the theme could be realized only working on a sample covering firms not listed in the Stock Market, that could bring out the complexity of control structures realized within a family circle or through syndicates. A comprehensive analysis is going to be the object of future work. Our results seems in line with Bianchi et al. (2008), due to the remarks on the recent evolution of our stock market. The pattern of ownership concentration could be inappropriate for newly listed firms, marked out by a simpler control structure. A sensitivity analysis could also show which are the bounds that allow keeping the control while lowering the ownership. This information could drive relevant policies both for private and public firms, and it will be the matter of future work.

Acknowledgments GR thanks Misako Takayasu, Hideki Takayasu, Tsutomu Watanabe and Nobuyasu Ito, and all other very kind APFA7 staff members, who made her stay in Japan possible through local grants. AMD and GR thank Mauro Costantini and Fabio Di Dio for their advices on econometrics matters. Last, but not least, AMD and GR thank the anonymous referees that gave fruitful hints for improving the paper.

References

Aganin A, Volpin P (2005) The history of corporate ownership in Italy. In: Morck Randall K (ed) The history of corporate governance around the world. University of Chicago Press, Chicago, pp 325–361

Bianchi M, Bianco M, Enriques L (2002) Pyramidal groups and the separation between ownership and control in Italy. In: Becht M, Barca F (eds) The Control of Corporate Europe, pp 154–188

Bianchi M et al (2008) The evolution of ownership and control structures in Italy in the last 15 years. In: Proceedings of the conference corporate governance in Italy: 10 years after the consolidated Law on Finance (TUF) organized by the Bank of Italy in Rome, 12/11/2008, http://www.bancaditalia.it/studiricerche/convegni/atti/corp_gov_it/session1/evolution_ownership_control_structures.pdf

Chapelle A, Szafarz A (2005) Controlling firms through the majority voting rule. Physica A 355:509–529

Dietzenbacher E, Temurshoev U (2008) Ownership relations in the presence of cross-shareholding. J Econ 95(3):189–212

Flath D (1992) Indirect shareholding within Japan's business groups. Econ Lett 38:223–227

Garlaschelli D, Battiston S, Castri MV, Servedio DP, Caldarelli G (2005) The scale-free topology of market investments. Physica A 350:491–499

Huber J, Ryll S (1989) Kapitalbeteiligungen von Bund und Laendern. Eine Analyse auf Basis einer Verflechtungsrechnung. Zeitschrift fur offentliche und gemeinwirtschatliche Unternehemen 12(3):287–305

Newman MEJ (2003) Mixing patterns in networks. Phys Rev E 67:026126

Newman M, Barabasi A-L, Watts DJ (2006) The structure and dynamics of networks. Princeton Studies in Complexity Princeton University Press, Princeton

Salvemini MT, Simeone B, Succi R (1995) Analisi del possesso integrato nei gruppi di imprese mediante grafi. L'Industria 4(XVI):641–662

Trusina A, Maslov S, Minnhagen P, Sneppen K (2004) Hierarchy measures in complex networks. Phys Rev Lett 92:178702

Turnovec F (1999) Privatization, ownership structure and transparency: how to measure the true involvement of the state. Eur J Polit Econ 15:605–618

Turnovec F (2005) Arithmetic of property rights: a leontief-type model of ownership structures. Working paper UK FSV-IES 70

Volpin P (2002) Governance with poor investor protection: evidence from top executive turnover in Italy. J Financ Econ 64:61–90