



# Evaluating risks-based communities of Mafia companies: a complex networks perspective

Nicola Giuseppe Castellano<sup>1</sup> · Roy Cerqueti<sup>2,3</sup> · Bruno Maria Franceschetti<sup>4</sup>

Accepted: 29 March 2021  
© The Author(s) 2021

## Abstract

This paper presents a data-driven complex network approach, to show similarities and differences—in terms of financial risks—between the companies involved in organized crime businesses and those who are not. At this aim, we construct and explore two networks under the assumption that highly connected companies hold similar financial risk profiles of large entity. Companies risk profiles are captured by a statistically consistent overall risk indicator, which is obtained by suitably aggregating four financial risk ratios. The community structures of the networks are analyzed under a statistical perspective, by implementing a rank-size analysis and by investigating the features of their distributions through entropic comparisons. The theoretical model is empirically validated through a high quality dataset of Italian companies. Results highlights remarkable differences between the considered sets of companies, with a higher heterogeneity and a general higher risk profiles in companies traceable back to a crime organization environment.

**Keywords** Companies financial risk indicator · Organized crime · Complex networks · Clustering coefficient · Entropy · Rank-size analysis

**JEL Classification** M19 · M49 · C18 · C63

## 1 Introduction

It is generally known that criminal organizations spread their wings over almost every kind of economic activity. In this respect, Italy represents a paradigmatic case (see the next section for a discussion on this point). In particular, companies in Italy can be effectively

---

✉ Nicola Giuseppe Castellano  
nicola.castellano@unipi.it

<sup>1</sup> Department of Economics and Management, University of Pisa, Via Ridolfi, 10, 56124 Pisa, Italy

<sup>2</sup> Department of Social and Economic Sciences, Sapienza University of Rome, Piazzale Aldo Moro, 5, 00185 Rome, Italy

<sup>3</sup> School of Business London, London South Bank University, London SE1 0AA, UK

<sup>4</sup> Department of Economics and Law, University of Macerata, Via Crescimbeni, 20, 62100 Macerata, Italy

clustered in the ones involved in illegal business and with strong connections with the Organized Crime—the so-called *Mafia* companies—and the other ones, the *No Mafia* companies. Mafia Companies can be owned directly by individuals affiliated to the Mafia, and are used as vehicles to legitimize an economic activity. As alternative, Mafia Companies can be owned by ‘collusive’ entrepreneurs, i.e. individuals that are interested to establish cooperative relationships with the Mafia, with the purpose to take some economic or personal advantages (Sciarrone 2009). Mafia Companies represent a research topic of peculiar interest, since they manipulate and monopolize financial markets, traditional institutions, and other legitimate industries (Federal Bureau of Investigation). They threaten many aspects of how humans live, work and do business. In short, these firms have a tendency to undermine democracy, the environment and the livelihoods and wellbeing of communities (Allum et al. 2019).

This paper aims at implementing a comparison between Mafia companies, considered as a whole, and No Mafia ones. The comparison is conducted with descriptive purposes by emphasizing in particular the differences between such groups expressed in terms of financial risk profile. Since existing knowledge about the investigated phenomenon is limited, the level of complexity and ambiguity of is undoubtedly relevant, also in consideration of the very wide number and nature of factors involved; therefore, a complex system research approach might be suitable, allowing to structure Mafia and No Mafia Companies data as networks. Consistent with Cinelli (2020), the term network refers to the structure of data analyzed through a complex networks theory approach, with the intent to extract new knowledge useful to contribute to the comprehension of the characteristics of Mafia Companies. In so doing, we are particularly close to Ozgul (2016) and Villani et al. (2019), where the authors describe illegal activities linked to terrorism—the former paper—and the Organized Crime—the latter one—by discussing the properties of such criminal networks through complex networks instruments. Interesting to mention also Grassi et al. (2019), where the authors identify the leaders in the Mafia organization through centrality measures of suitably constructed complex networks. In the same line, Mastrobuoni and Patacchini (2012) discuss the topological structure of the Mafia network by implementing an extensive complex networks-based analysis, and find evidence of a hierarchical structure of such illegal organizations.

However, the quoted papers deal with individuals. Rather than this, this paper explores companies and their financial characteristics. Specifically, we here discuss the risk profiles of the individual companies and of the overall set generated by them.

We assume that the companies of the considered groups are mutually interconnected. Such interconnections are driven by their similarity—to be intended here in terms of risk—so that a strong link between two companies captures the fact that they share a common financial risk profile whose entity is large. This assumption is in line with several studies, which provide evidence that financial indicators based on companies’ financial structure show some commonalities among companies that are neutral to criminal organizations relatively to those who are not (see e.g. Ravenda et al. 2015b; Fabrizi et al. 2017; La Rosa et al. 2018). Moreover, such a condition is in agreement also with the well-known empirical evidence that financial variables are more correlated when the market is in a high-volatility phase (see e.g. Bartram and Wang 2005; Forbes and Rigobon 2002; Ramchand and Susmel 1998; references therein contained). Under a purely methodological viewpoint, similarity-based interconnections represent the ground of the well-established context of *homophily*, i.e. the attitude of two nodes sharing similar characteristics to be strongly linked. Relevant examples in the literature can be found in Cinelli et al. (2016), Mollgaard et al. (2016). Importantly, homophily is not only associated to the attributes of the nodes in

a social networks context, but can also be related to the segregation of similar entities in a financial and economic framework. This is the proposal of Elliott et al. (2014), where the authors presents a homophily-based model for countries and industrial sectors. In our setting, we adopt the financial perspective of this relevant paper.

The different sets of Mafia and No Mafia companies are assumed to form two different networks whose weighted arcs are constructed according to the same rule. A complex network approach aimed at describing similarities has also previously used by Cinelli et al. (2016), Cerqueti et al. (2018), D'Arcangelis et al. (2019) and D'Arcangelis et al. (2020), only to name a few.

The financial risk profile is assessed through the construction of a risk indicator, which aggregates four ratios derived by financial analysis, which are significant in expressing to what extent different stakeholders (i.e. suppliers, workforce, financial lenders and the overall community of creditors) are exposed to a financial risk deriving by their relation with a given company. The four indicators are the following: (a) Days Payable Outstanding (DPO, hereinafter); (b) Timely Payments to Social Institutions (TPSI, hereinafter); (c) Funded Capital Ratio (FCR, hereinafter); (d) Long Term Financial Debt Coverage (LTFDC, hereinafter). The overall risk indicator is defined as a fair sum of the four financial ratios, so that it is conceptualized by assigning an equal weight to its constitutive parameters. In so doing, we are in line with literature on corporate disclosure, where the adoption of self-constructed indexes is quite common. Unweighted indexes are generally preferred since they are easier to calculate and allow to reduce the subjectivity bias (see e.g. Cooke 1989; Ahmed and Courtis 1999).

Thus, in the context of risk and of our definition of the connection between two companies, we compare and provide a deep exploration of the community structures of the considered networks. At this aim, we employ the clustering coefficient of the nodes of the networks, which is a statistical measure for complex networks particularly suitable for our purposes. In this respect, we refer to the conceptualizations of the clustering coefficients for weighted and unweighted, directed and undirected networks in Barrat et al. (2004), Cerqueti et al. (2020), Clemente and Grassi (2018), Fagiolo (2007), Onnela et al. (2005) and Watts and Strogatz (1998). In our specific context, we adopt the definition of clustering coefficient of Onnela et al. (2005), which is particularly suitable in our framework of weighted and undirected networks.

The exploration of the community structures of the networks allows to gain insights on the strength of the connections between the individual companies in the Mafia and No Mafia cases, along with their tendency to form clusters of highly risky and interconnected companies. It is worth mentioning, however, that problems of interpretability caused by the rigidity of the network structure may arise when comparing the networks (Cinelli et al. 2017, 2020).

To pursue our purpose, we investigate the empirical distribution of the clustering coefficients under two very different perspectives. By one side, we analyze the Shannon entropies (introduced by Shannon 1948) of the considered samples and of a large set of sub-samples of highest and lowest clustering coefficients. In so doing, we are able to discuss the deviations of the risk profile-based community structures associated to the companies from the very relevant cases of companies absolute homogeneity (i.e., uniform distribution, which is associated to the maximum level of entropy) or absolute heterogeneity (i.e., concentration over a unique clustering coefficient which appear when entropy reaches its minimum value). In order to have an intuitive view of the role of entropy, it is worth mentioning the applications of such an instrument in portfolio theory (see e.g. Bera and Park 2008; Mercurio et al. 2020; Pola 2016). In such a context, a large value of the entropy of

the weights of a portfolio is associated to a high level of portfolio diversification. Thus, the employment of Shannon entropy is quite informative in our framework, in that it allows to state whether the risk profiles of Mafia and No Mafia companies follow scattered or common patterns. The usefulness of Shannon entropy is witnessed by its popularity in a wide strand of applied science literature, including of course economics and management (see e.g. the recent contributions of Bartolacci et al. 2015; Chao et al. 2015; Fedajev et al. 2020; Karagiannis and Karagiannis 2020; Rosser 2016; Yang 2018).

By the other side, we apply a rank-size analysis of the nodes by taking as size the level of clustering coefficient. Companies are ranked in decreasing order according to the values of the clustering coefficients, so that the highest value of such a coefficient is associated to unitary rank. Rank-size analysis allows to derive a best fitting curve-based systemic view of the considered networks. In details, the investigation and comparison of the calibrated curve parameters and of the goodness of fit indicators can be effectively used to gain relevant information on the discrepancies between Mafia and No Mafia companies, and of their risk profiles in a systemic contextualization. Rank-size analysis is one of the most popular methodological statistical investigation device when one aims at deriving an overall view of a phenomenon on the basis of some observations, by taking under special consideration how a quantitative characteristics of the phenomenon drives the collocation of the individual observations in terms of rank. Such a methodology is able to translate a set of points—clustering coefficients, in our case—to an unified system having clear macroscopic characteristics. Moreover, the rank-size analysis is also able to provide insights on the mechanism of random growth generating the explored system (see Gabaix 2009 for a detailed explanation of this aspect). Even if the study of such a mechanism is well-beyond the scopes of the present paper, the possibility of discussing the generator of the ranked data is a further motivations for analyzing Mafia and No Mafia companies under the rank-size perspective. Thus, it is not unexpected that the literature on this field is rather wide, and includes authoritative contributions in the contexts of management and economics (see e.g. Ausloos and Cerqueti 2016; Bartolacci et al. 2019; Cerqueti and Ausloos 2015a, b; Briant et al. 2010; Gabaix 1999a, b). In the context of rank-size analysis, we adopt a best fit curve of third degree polynomial type. Such a choice has been driven by a preliminary visual inspection of the scatter plot of the rank-size data, but also by the intuitive interpretation of the calibrated parameters of this specific function. Moreover, the discrepancy between the best fit curve and the scatter plot—to be intended under a statistical point when looking at the goodness of fit parameters or under a purely visualization perspective—allows to identify the presence of some outliers for both networks of Mafia and No Mafia companies. This is a well-known property of this type of analysis, called *king and vice-roys effect* (see e.g. Laherrere and Sornette 1998; Ausloos 2013), being the highest outlier the king and the other ones the vice-roys.

In the next section we will present the formal statement of our main research hypothesis of our study, along with a supporting literature review. However, we can preannounce that results suggest that Mafia companies are more heterogeneous than No Mafia ones, especially when referring to the subsample of high values of the clustering coefficients. Basically, the financial risks of Mafia companies is higher than those of No Mafia ones, and the former set of companies is associated to stronger communities of clusters of highly risky companies.

The rest of the paper is organized as follows. Section 2 contains a brief literature review on the applied topic we deal with, to support the investigated research hypothesis—which is formally stated at the end of the section. Section 3 outlines the details of the theoretical weighted network model. In such a section, the overall risk indicator for the companies is

also introduced. Section 4 contains the description and a deep exploration of the empirical dataset used for the comparison experiment; moreover, it presents and discusses the adopted methodological tools. Section 5 is devoted to the illustration of the results and to the discussion of them. Last section offers some conclusive remarks.

## 2 Literature review and statement of the research hypothesis

As already mentioned above, a proper description of the Italian productive structure cannot avoid an explicit reference to the role of the Organized Crime in the economic environment (see e.g. Allum et al. 2019; Esposito et al. 2019; Savona and Riccardi 2018; Savona and Calderoni 2016; Pinotti 2015; Calderoni 2011; Paoli 2004; Arlacchi and Ryle 1986) since the global revenues generated by Italian mobsters reached 211 bln Euros in 2017, which accounts approximately for 12% of Italy's GDP (ISTAT, 2019). However, ISTAT's estimates refer to non-observed economy, which includes underground, illegal and other shadow productive activities but little attention is given to the instrumentality used legally by criminal organizations to implement their business ventures, i.e. companies that are legally registered in the Italian Business Register but traceable back to a Mafia environment.

In this respect, it is worth noting that Mafia Companies can be used for different purposes and hold different characteristics accordingly. In particular, within the general definition of Mafia companies three sub-categories can be identified, namely: screen companies; papermakers and star companies (La Rosa and Paternostro 2015). Screens and papermakers, are generally used by criminal organizations for money laundering services or for supporting activities of other kind; their financial statements may show some anomalies such as: no revenues in combination with operating costs; or substantial balance between revenue and costs; significant fluctuations in revenues over time; abnormal liquid funds, at least if compared with the volume of operating activities (Fabrizi et al. 2017). Conversely, star companies effectively do business in a given industry, and in several cases they show stable positive economic and financial performance. Star companies may be the mean for organized crime to create stable relations with institutions, governmental agencies, politics and create the suitable environment where bribery initiatives might take place.

In general, prior studies suggest that Mafia companies are more pervasive in some specific industries since the industry in which a firm operates is an essential variable in explaining illegal activity (Daboub et al. 1995). Firms in specific industries are more likely to commit illegal acts (Baucus and Near 1991; Simpson 1986) and show similar rates of illegal activity (Cressey 1976). Mafias demand to exert power over specific business sectors (Allum et al. 2019) and date sets of prior studies confirm that claim. For example, Ravenda et al. (2015a, 2015b) used a sample of mafia-related firms mainly operating in building and construction, wholesale and retail trade, transportation and storage industries. Similarly, Savona and Riccardi (2018) and Savona and Berlusconi (2015) mapped a wide variety of business sectors in which mobsters play a vital role and conclude that construction, wholesale, transportation, healthcare and waste management are involved mostly. The same industries are depicted in our data set as well. Apart from those industries stimulated by the possibility of intercepting substantial flows of public resources such as waste management, healthcare, and construction bids, it is undeniable that clans are investing in other areas, such as utilities, hotels and restaurants, real estate, agriculture, import/export, and financial services (Transcrime 2013; Savona and Berlusconi 2015). However, both

academic and practitioner literature seem to support homogeneity in terms of industry. Moreover, Article 53 of the Italian Law No. 190 of November 6, 2012, on the prevention and repression of corruption identifies the most vulnerable economic sectors to mafia infiltration, i.e., water supply, sewerage, waste management and remediation activities (EU NACE code E), construction (EU NACE code F), wholesale and retail trade (EU NACE code G), transportation and storage (EU NACE code H). Thus, it is reasonable to assume homogeneity in terms of industry.

In the present paper, we explore the deviation between Mafia companies and No Mafia ones by adopting a financial risk perspective. In so doing, we contribute to the macroarea of corporate risk (see Cao et al. 2015; Firth and Smith 1995; Koutmos et al. 2018, only to name a few). We contextualize Mafia and No Mafia companies in this field of studies; this might be helpful, to some extent, in explaining uncertainty in financial systems.

In the literature, the topic of Mafia companies has not attracted a great number of scholars, although relevant negative effect may derive on economic growth at local level, particularly in those areas where criminal organizations are remarkably active and concentrated (Mirenda et al. 2019). To the best of our knowledge the financial risk perspective associated to Mafia companies has not been investigated explicitly yet; however, extant studies provide evidences that motivate both our research purpose, as well as the adopted approach. In particular, in the literature on criminal activities, some studies show that mafia companies hold similarities in their financial risk profiles (represented through financial indicators), which express higher risk levels compared to companies not belonging to criminal organizations (Fabrizi et al. 2017; Ravenda et al. 2015b). Mirenda et al. (2019) show that companies which over time experience infiltrations by criminal organizations are characterized by a progressive deterioration in the leverage ratio and consequently in their financial stability, while in the years that precede the infiltration a flat trend was occurring. Going more in depth in the composition of the level of indebtedness, it is worth noting that financial debts in Mafia companies show generally a low incidence on total assets (Di Bono et al. 2015); consequently, a great percentage of debts must be of operating nature (i.e. payables to suppliers, to the employees, to the tax authority or to social security institutions). Drawing on the above mentioned references, in our analysis of exploratory kind, we use financial ratios related to financial indebtedness referred to specific categories of creditors (namely suppliers, banks and the company workforce) and we formulate the following research hypothesis:

**H1** *The risk profile of the system of Mafia companies is higher than that of the No Mafia ones.*

### 3 Network model

We build two different networks of companies: the first one is associated to the Mafia companies, the other one collects the other (No Mafia) companies. The distinction is based on the inner characteristics of the considered companies. It will be clear in the next Section, where the empirical dataset will be described in details.

We collect the Mafia companies in a set  $V^m$  and the No Mafia ones in a set  $V^{nm}$ . Such sets represent the set of the nodes of the two networks under consideration.

Nodes are connected according to a common rule. To describe such interconnections, we refer hereafter to the generic set of nodes  $V$ . In particular, we consider weighted

connections, so that the generic link connecting the nodes  $i$  and  $j$  in  $V$ —namely,  $(i, j)$ —is associated to a nonnegative weight  $w_{ij}$  which represents the strength of the connection between  $i$  and  $j$ . Such weights are listed in a squared matrix  $\mathbf{W} = (w_{ij})_{i,j \in V}$  whose order is the cardinality of  $V$ —namely,  $|V|$ . Matrix  $\mathbf{W}$  is the weighted adjacency matrix of the network.

We here face the risk profile of the considered companies. Thus, we first need to define the risk profile of a generic company  $k \in V$ .

We derive risk indicators by financial statement analysis. All the indicators are based on company liabilities, which reasonably may be considered as relevant driver of risk in the perspective of the several categories of creditors. On the basis of data availability (see the next section on this point) and as preannounced in the Introduction, we here consider four of them: (a) DPO; (b) TPSI; (c) FCR; (d) LTFDC; for all the four indicators, the higher the values, the higher the financial risk profile associated to the company (see the next section on these points).

We denote them when associated to a generic company  $i \in V$  by  $x_a(i), x_b(i), x_c(i), x_d(i)$ , respectively.

The aggregated indicator of the  $i$ -th company is then

$$X_i = x_a(i) + x_b(i) + x_c(i) + x_d(i), \quad \forall i \in V. \quad (1)$$

The indicator in (1) is conceptualized to treat in a fair way the terms composing it, without assigning prominence to any risk parameter coming from the balance sheet. In so doing, we are in line with literature on corporate disclosure, where the adoption of self-constructed indexes is quite common (see e.g. Castellano et al. 2019). Generally, unweighted indexes are preferred in this field of research, since they are easier to calculate and allow to reduce the subjectivity bias (see e.g. Cooke 1989; Ahmed and Curtis 1999), whereas weighted indexes require personal judgements about the weights to be assigned to every item. We are now ready to introduce the way in which two companies are connected in the network.

The research approach used in this paper is quite common in research on complex networks. *Homophily* is the term used to address the link between two nodes which are similar according to a set of pre-determined characteristics and being this true means that the number of links is able to represent how spread are nodes holding similar characteristics (see e.g. Cinelli et al. 2016; Mollgaard, 2016). In this line, we assume that two companies  $i, j \in V$  exhibits a high level of interconnection—i.e., a high value of  $w_{ij}$ —when they have similar risk profiles and, at the same time, such risk profiles are large—according to the definition of risk profile given in (1). We also assume that very small levels of risk profiles are associated to low connections, even if such levels are similar. Moreover, in order to focus specifically to the community structures generated by the mutual interconnections of different nodes, we conveniently assume that self-connections are not allowed.

Thus, we define the weights  $w$ 's as follows:

$$w_{ij} = \begin{cases} \frac{X_i + X_j}{2 \cdot 400 \cdot (|X_i - X_j| + 1)}, & \text{if } i \neq j; \\ 0, & \text{if } i = j. \end{cases} \quad (2)$$

By definition, one can notice that  $w_{ij}$  increases with respect to  $X_i$  and  $X_j$  and decreases with respect to  $|X_i - X_j|$ . Moreover, the range of the weights in (2) is  $[0, 1]$ , with  $w_{ij} = 1$  when  $X_i = X_j = 400$  and  $w_{ij} = 0$  when  $X_i = X_j = 0$ .

We denote the networks of the Mafia companies and No Mafia ones by  $N^m = (V^m, \mathbf{W}^m)$  and  $N^{nm} = (V^{nm}, \mathbf{W}^{nm})$ , respectively.

## 4 Data and methodology

This section is devoted to the description of the considered dataset and the methodological devices used to explore and compare the considered networks.

### 4.1 Description of the dataset of the considered companies

Under the Italian Criminal law (Royal Decree n. 1398 of 19 October 1938) after the first instance of court confiscation Mafia companies are assigned to the Italian agency for the management of seized and confiscated assets, namely Agenzia Nazionale Beni Sequestrati e Confiscati (hereinafter ANBSC). Therefore, confiscated companies belonging to the list provided by ANBSC, where chosen.

The initial population comprised all 1705 confiscated companies. In Italy, Limited liability companies and cooperatives have to deliver their annual reports to the official business register of the Italian Chambers of Commerce while partnerships, sole proprietorships and other legal forms (e.g. associations, trusts, etc.) are not required to file annual reports. Accordingly, 391 (22.93%) partnerships, 408 (23.93%) sole proprietorships and 59 (3.46%) other legal forms, with no available financial data were removed because of different normative environments. The remaining 847 companies are limited liability companies (830) (48.68%) and cooperatives (17) (1%).

Firms provided by ANBSC have all been confiscated by final judgment but unfortunately most of them are not traceable due to secrecy reasons. Only 172 (10%) companies could be identified either by name or by value-added tax number—i.e. VAT. Thus, another set of 217 firms found in AIDA, the Italian Bureau Van Dijk database, with status confiscated has been added to the initial group of Mafia companies. The financial statements for all firms are obtained from AIDA, database but their small size, their inactivity or their liquidation status means that only 231 out of 389 Mafia companies, have financial statements available on AIDA. For each company, we reviewed the first recent financial statement available prior to the confiscation year as once confiscated Mafia companies may lose their distinctive characteristics (Ravenda et al. 2015a). In doing so, we have obtained companies' financial statements ranging from 2002 to 2017. Moreover, some missing financial statement items on AIDA data for the calculation of selected variables (i.e. DPO, TPSI, FCR and LTFDC, see the next Subsection for their description) in some years further reduced the number of Mafia companies which ends up being 97. The remaining 97 companies were further classified by EU NACE Rev. 2 codes (16 manufacturing; 13 water supply, sewerage, waste management and remediation activities; 13 construction; 19 wholesale and retail trade, repair of motor vehicles and motorcycles; nine transportation and storage; three information and communication; nine real estate; six administrative and support service activities; and nine arts, entertainment and recreation companies).



**Table 1** Formulas for the computation of the financial ratios associated to companies risks, along with their variation ranges

Ratio	Formula	Range
DPO	Accounts payable $\times$ 360/purchases with delayed payments (+ VAT*)	(0, $+\infty$ )
TPSI	Debts for social contributions/max Debt in case of timely payments**	(0, $+\infty$ )
FCR	Equity + long term loans /fixed assets	(0, $+\infty$ )
LTFDC	Long term financial loans /property	(0, $+\infty$ )

\*VAT stands for value-added tax, as mentioned above. \*\* Max debt in case of timely payments is the maximum theoretical value of debt to social security and welfare agencies and is computed as the ratio between Social Security Expenses and 12

Firms were also grouped by number of employees expressed in annual work units (AWU) and by total assets (TA) to be able to decide whether the companies are micro (53 companies), small (22 companies), or medium-sized (22 companies)<sup>1</sup>. Due to industrial classification and size differences, these groups are not completely homogeneous. A careful selection of No Mafia companies was attempted to construct the control group. Companies have been matched by regional headquarter first. This means that we looked for No Mafia companies in the same Italian region in which, each analyzed Mafia company has its own headquarter. However, the probability to choose a non-confiscated Mafia company is very high in some Italian regions (e.g. Sicily, Calabria, Apulia, and Campania) thus only companies found on the regional White lists produced by the Ministry of Interior have been considered. These lists<sup>2</sup> contain detailed information about companies that voluntarily have accepted on-going and in-depth checks, performed by the Italian authorities, aimed at preventing criminal infiltrations in the specific firm. Therefore, companies on the list are assumed to be No Mafia companies.

We further paired No Mafia companies by industry, number of employees, and total assets. Lastly, we analyzed for each No Mafia company the same firm year observation as we did for Mafia companies. The final control group consists of a set of 127 No Mafia companies.

## 4.2 The risk indicators of the companies

According to formula (1), the overall risk indicator of the companies  $X$  is created by summing four financial ratios, namely  $x_a, x_b, x_c, x_d$ , which may express the financial risk of a company under four different perspectives. Table 1 summarizes the formulas for the computation of the ratios—which are labeled by their acronyms—along with their variation ranges.

The four indicators have been derived by building on previous literature on criminal organizations. Mafia companies show higher level of debts on total assets relative to No Mafia ones (Ravenda et al. 2015b; Fabrizi et al. 2017), but at the same time Mafia Companies shows a lower level of bank debt (La Rosa et al. 2018; Ravenda et al. 2015b), which

<sup>1</sup> Since Mafia companies have the power to discourage competition their income statement may report biased revenue volumes. Therefore, we decided to exclude the annual turnover parameter recommended by EU recommendation 2003/361.

<sup>2</sup> Lists are official, and are available online at <http://www.prefettura.it/>.

**Table 2** Main descriptive statistics for the financial risk ratios. Case of Mafia companies

	DPO	TPSI	FCR	LTFDC
Mean	156.24	2.02	2.24	2.57
SD	90.29	0.80	2.05	1.92
Mean/SD	1.73	2.52	1.10	1.34
Min	4.67	0.79	0.04	0.20
Max	373.75	3.87	5.85	6.90
Q1	90.26	1.45	0.53	1.15
Median	136.94	1.82	1.09	1.88
Q3	192.96	2.57	4.44	3.63
Skewness	0.91	0.48	0.65	1.10
Kurtosis	0.18	-0.60	-1.23	-0.07

may reasonably be interpreted as a major inclination for Mafia companies to cover their financial needs by resorting to operating debts, such as accounts payables. By the way, accounts payables represent sometimes the way used to let dirty money enter into the Mafia company for money laundering purposes. Coherently with the cited literature we use the DPO as first indicator of risk, measuring the days needed on average by a company to settle its accounts with suppliers.

According to Ravenda et al. (2015a, 2015b), indicators based on labor and tax evasion might also be meaningful in the explanation of differences between Mafia and No Mafia companies. In line with this, Arlacchi (2007) affirms that among their “competitive advantages”, companies belonging to criminal organizations frequently are not compliant with payments on social contributions; therefore the second ratio—i.e., TPSI—provides a measure of financial risk in the perspective of a company workforce. When TPSI is equal to 1 (or lower) the company is respecting (or anticipating) the terms of payment fixed by the law for social and employee contributions, while the higher is the ratio, the higher is the amount of money unpaid. The third and fourth ratio derive by considerations about the adequacy of the financial structure of a company (Arlacchi 2007; Dupla et al. 2012; Ravenda et al. 2015a, b, only to name a few). The FCR is a measure of financial stability: when the ratio is higher than 1, the financial structure of a company is in equilibrium and the amount of long-term debts and equity is compatible with the long term financial needs generated by the fixed assets, and this should prevent a liquidity crisis. On the other hand, too high values of the ratio may not necessarily be a positive sign, since the permanent liabilities in excess could be used as example to create liquidity funds available for activities instrumental for the criminal organization. The fourth ratio, the LTFDC, is similar to the previous one but focused on creditors of financial nature. Usually property is requested as guarantee when opening a mortgage or other kinds of long term financial debts; so when the amount of debts overwhelms the value of property, this could be the sign that debt is obtained through corruption or other persuasive methods. Data needed to calculate the four indicators are included in the mandatory information of financial statements.

It is worth noting that measures of profitability, which is a relevant driver of companies’ financial stability, could also have been included in the financial risk index. However, as already mentioned in Sect. 2, the different subcategories of Mafia companies (screen, papermakers and star) may show remarkable differences in revenues and operating costs (Fabrizi et al. 2017; La Rosa and Paternostro 2015). Therefore, Mafia companies belonging to such categories cannot be always reasonably compared with the No Mafia ones in terms

**Table 3** Main descriptive statistics for the financial risk ratios. Case of No Mafia companies

	DPO	TPSI	FCR	LTFDC
Mean	153.09	0.7	1.2	1.09
SD	94.71	0.44	1.07	0.82
Mean/SD	1.62	1.6	1.11	1.32
Min	2.38	0	0	0
Max	383.36	1.9	4.43	3.09
Q1	75.43	0.34	0.45	0.5
Median	134.97	0.64	0.87	0.82
Q3	229.42	0.95	1.72	1.76
Skewness	0.33	0.75	1.31	0.95
Kurtosis	-0.85	0.06	1.26	-0.07

**Table 4** Main descriptive statistics for the overall risk index

	Mafia companies	No Mafia companies	All companies
Mean	168.24	93.93	126.11
SD	81.56	41.24	72.12
Mean/SD	2.06	2.28	1.75
Min	54.91	11.5	11.5
Max	375.98	203.16	375.98
Q1	101.78	61.42	76.01
Median	160.74	84.81	110.85
Q3	203.89	122.74	160.92
Skewness	0.85	0.44	1.34
Kurtosis	-0.06	-0.41	1.85

of profitability. Since we do not control the type of Mafia company in this first exploratory study, we decided not to include profitability measures in the financial index.

Descriptive statistics about the four financial indicators for Mafia companies and No Mafia ones are provided in Tables 2 and 3, respectively.

The distributions of data for Mafia companies and No Mafia ones look substantially similar, since they show same levels of data dispersion, as well as skewness and Kurtosis. The distributions show a slight positive skewness, which is reasonable to be expected, considering that all the ratios have a finite minimum value but an infinite maximum. Kurtosis is close to zero in the majority of the cases, showing the existence of rather mesokurtic distributions. Of course Mafia companies' distributions are all shifted around higher mean values, these being remarkably higher for TPSI, FCR and LTFDC, thus providing evidence that higher financial risk levels may be associated to such category of companies. The observations about the four risk ratios have been normalized in  $[0, 100]$  in order to calculate the overall risk index  $X$ , by using the following simple procedure:

$$x_{\star}(i) \mapsto \frac{x_{\star}(i) - \min_{j \in V} x_{\star}(j)}{\max_{j \in V} x_{\star}(j) - \min_{j \in V} x_{\star}(j)} \times 100, \quad \star = a, b, c, d. \quad (3)$$

**Table 5** Standardized Cronbach’s Alpha for the overall risk indicator  $X$  when one of the four financial risk ratios is set to zero in formula (1)

	DPO	TPSI	FCR	LTFDC
Cronbach’s Alpha	0.63	0.59	0.53	0.51

In the columns, the removed financial risk ratio, as expressed through its acronym for an easy interpretation

To avoid cumbersome notation, we will refer hereafter to the  $x$ ’s directly as the normalized variables, according to formula (3). After standardization, data range from 0 to 100 in all the four indexes. The overall risk index  $X$  has been calculated as unweighted sum of each risk component. Descriptive statistics about the overall risk index, calculated for Mafia companies and No Mafia ones, are shown in Table 4.

The average risk score for Mafia companies and No Mafia ones compared to the total mean is higher and lower respectively, supporting our expectation that Mafia companies may be considered of relatively higher risk. Reasonably, the comments made for descriptive statistics of the four risk components, generally also apply to the overall risk index distribution. It is worth noting that the total data distribution (i.e., the distribution of the entire set of companies) shows a more pronounced positive skewness and slightly leptokurtic shape, whereas the distributions of Mafia and No Mafia companies are mesokurtic. To validate the overall risk indicator, we have calculated the Standardized Cronbach’s Alpha in order to test the internal consistency of the risk items included, and we obtained an overall value of 0.7, which according to the general rule of thumb used to interpret the indicator, is the lower threshold needed to consider the overall risk indicator  $X$  as a reliable measure of risk. In this respect, we have also computed the Standardized Cronbach’s Alpha to a modified version of the indicator  $X$  when, alternatively, one of the four financial risk ratios is set to zero in formula (1). Results are reported in Table 5.

Table 5 shows that the Standardized Alpha decreases in all cases and becomes less than 0.7 when removing single items. This represents a supporting argument for selecting not less than four financial risk ratios for the construction of the overall risk indicator, hence giving consistency to our approach of defining  $X$  in (1) as a corporate financial risk measure.

### 4.3 Methods of investigation

The basis of the analysis is the assessment of the community structure of the nodes of the networks  $N^m$  and  $N^{nm}$ . To this end, we first compute the clustering coefficient of the generic node  $i \in V$ . According to the presented framework, we employ a weighted version of the clustering coefficient of  $i$ —namely,  $c_i$ —introduced by Onnela et al. (2005) and given by

$$\tilde{c}_i = \frac{\sum_{j,k \in V} (w_{ij}^{1/3} w_{ik}^{1/3} w_{jk}^{1/3})}{(|V| - 1)(|V| - 2)}. \tag{4}$$

We denote by  $\mathbf{c} = (c_1, \dots, c_{|V|})$  the vector of the clustering coefficients in (4). As usual, we will refer to superscript  $m$  and  $nm$  to denote the clustering coefficients of the nodes of  $N^m$  and  $N^{nm}$ , respectively.

### 4.3.1 Entropic comparison of the empirical distributions

We here aim at providing a comparison of the empirical distributions of the elements of  $\mathbf{c}^m$  and  $\mathbf{c}^{nm}$ .

In particular, we analyze the statistical disorder of the clustering coefficients in terms of Shannon entropy. In so doing, we are able to assess the distance between the considered distributions and the uniform one (case of maximal entropy equals to  $\log(|V|)$ ) or the Dirac-type distribution with probability concentrated over only one value (case of minimal entropy equals to zero). This information is of particular interest in our context, and we refer to the next section for a list and discussion of the obtained results.

Given the vector  $\mathbf{c} = (c_1, \dots, c_{|V|})$  of the clustering coefficients, we define the Shannon entropy of  $\mathbf{c}$  by

$$\mathcal{E}(\mathbf{c}) = - \sum_{k=1}^{|V|} \frac{c_k}{\sum_{i=1}^{|V|} c_i} \cdot \log \left( \frac{c_k}{\sum_{i=1}^{|V|} c_i} \right). \tag{5}$$

Beyond the original sample of clustering coefficients, we also compute the entropy of the subsamples associated to low and high levels of communities. At this aim, we consider the order statistics of the clustering coefficients  $\mathbf{c}^{ord} = (c_{(1)}, \dots, c_{(|V|)})$ , which is obtained by permuting the indexes in order to have  $c_{(1)} \geq \dots \geq c_{(|V|)}$ .

For the sake of simplicity and without being restrictive, we assume that  $c_k = c_{(k)}$ , for each  $k \in V$ , so that  $\mathbf{c}^{ord} = \mathbf{c}$  and one has

$$c_1 \geq \dots \geq c_{|V|}. \tag{6}$$

Consider now  $\alpha, \beta \in [0, 1]$ . We define the vectors of the  $\alpha$ -high and  $\beta$ -low communities—and denote them by  $\mathbf{c}_{H(\alpha)}$  and  $\mathbf{c}_{L(\beta)}$ —as follows:

$$\mathbf{c}_{H(\alpha)} = (c_1, \dots, c_{H(\alpha)}), \quad \mathbf{c}_{L(\beta)} = (c_{L(\beta)}, \dots, c_{|V|}),$$

where  $H(\alpha)$  and  $L(\beta)$  are integers in  $\{1, \dots, |V|\}$  such that  $c_{H(\alpha)}$  is the  $1 - \alpha$ -th percentile and  $c_{L(\beta)}$  is the  $\beta$ -th one of the distribution of the components of vector  $\mathbf{c}$ .

In the next Section, we will compute the entropies  $\mathcal{E}(\mathbf{c}_{H(\alpha)})$  and  $\mathcal{E}(\mathbf{c}_{L(\beta)})$  according to formula (5) and for a number of levels  $\alpha$  and  $\beta$ , for both cases of Mafia and No Mafia networks.

### 4.3.2 Rank-size analysis

The companies of the two networks are sorted in decreasing order in terms of their clustering coefficients, so that condition (6) is assumed to be satisfied. Thus, we assign rank  $r = 1$  to  $c_1$  and rank  $r = |V|$  to  $c_{|V|}$ . The clustering coefficients represent the size, and each size  $z$  is associated to a rank  $r$ .

Then, we implement a best fit procedure for assessing the shape of the curve representing the considered sample. A visual inspection of the scatter plot for both cases of  $N^m$  and  $N^{nm}$  suggests that a simple polynomial function seems to be suitable for satisfactorily fitting the data. Thus, we proceed by trying a best curve of polynomial type. In details, the clustering coefficient  $z$  is tried to be approximated by a polynomial type function of the rank. In so doing, the best fitting curve represents well an overall system which is built on the basis of an original sample.

**Table 6** Values of the entropies for the Mafia companies

	Entropy	Sample dimension	Max entropy	Entropy / Max entropy
all sample	4.084488475	97	4.574710979	0.89284077
$\alpha = 90$	1.106502017	9	2.197224577	0.50359077
$\alpha = 80$	1.85153195	18	2.890371758	0.640586092
$\alpha = 50$	3.372955488	49	3.891820298	0.866678117
$\beta = 50$	0.711532987	48	3.871201011	0.183801612
$\beta = 20$	0.153286845	18	2.890371758	0.053033609
$\beta = 10$	0.039251099	9	2.197224577	0.017863945

The first column contains the considered scenarios. Second column is devoted to the computation of the entropies, according to formula (5). Third column lists the cardinality of the considered sample, according to the individual scenarios. Fourth column contains the theoretical maximum entropy levels, computed as  $\ln(\text{Sample dimension})$ . The ratio of the values of the entropies with the maximum possible ones are also reported, for an easy interpretation of the results (see the last column)

**Table 7** Values of the entropies for the No Mafia companies. For a detailed description of the content of the Table, please refer to the caption of Table 6

	Entropy	Sample dimension	Max entropy	Entropy / Max entropy
all sample	1.597756233	127	4.844187086	0.329829588
$\alpha = 90$	0.357298529	12	2.48490665	0.143787506
$\alpha = 80$	0.627934639	24	3.17805383	0.197584645
$\alpha = 50$	1.25861281	64	4.158883083	0.30263241
$\beta = 50$	0.42790869	63	4.143134726	0.103281384
$\beta = 20$	0.052829336	24	3.17805383	0.016623172
$\beta = 10$	0.007066817	12	2.48490665	0.002843896

In the analysis, as we will see in the next Section, we successfully try a third-degree polynomial function of the type

$$z = f(r) = ar^3 + br^2 + cr + d, \quad (7)$$

where  $a, b, c, d \in \mathbb{R}$  are parameters whose values have to be calibrated.

The calibrated parameters have very specific meanings, which will be useful for the comparison of  $N^m$  and  $N^{mm}$  (see the next Section for comments on this point).

The best fit procedure leads to identification to a *king and vice-roys effect*—as defined by Laherrere and Sornette (1998)—which points to the presence of a remarkable outlier (the king) and some others less noticeable (the vice-roys) at high ranks  $r = 1, 2, \dots$ , leading to a detrimental effect on the goodness of fit parameters and on the visual appeal of the curve when compared to the scatter plot.

## 5 Results and discussion

This Section collects the results of the analysis and provides some related comments.

### 5.1 Entropy-based analysis

We here present the analysis of the distributions of the clustering coefficients in the cases of Mafia and No Mafia companies.

Different scenarios are treated: the entire sample of 97 Mafia companies and 127 No Mafia ones, but also  $\alpha = 50, 80, 90$  and  $\beta = 10, 20, 50$ . In so doing, we get insights on the left and right tails of the distributions of the clustering coefficients in terms of their distance with the cases of maximum entropy—which points to the highest level of randomness, represented by the uniform distribution case—and minimum entropy—which is associated to the deterministic case of concentration of the distribution in one value.

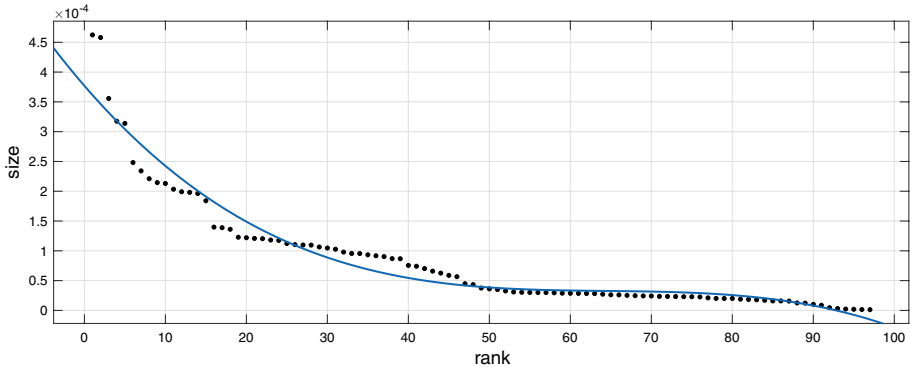
Tables 6 and 7 collect the outcomes of the analysis.

In the context of Mafia companies, Table 6 suggests that the distribution of the clustering coefficients is particularly close to the uniform distribution when the entire sample is considered. Such an outcome gives that companies are quite scattered in terms of their property of connecting with the others, when connections are driven by a similarity in their risk profiles. Substantially, the network of Mafia companies is rather heterogeneous in terms of their risk profile. The situation is quite different when No Mafia companies are considered. Indeed, Table 7 highlights that the entropy of the all sample case is quite low, with a remarkable distance from the maximum possible level of entropy. This means that clustering coefficients are rather concentrated around some specific values, and the distribution of companies in terms of their risk-based community structure levels is quite narrow. No Mafia companies appear to be quite homogeneous, in the sense that they can be viewed as belonging to a unique cluster with similar interconnections structures and, thus, similar risk profiles characteristics.

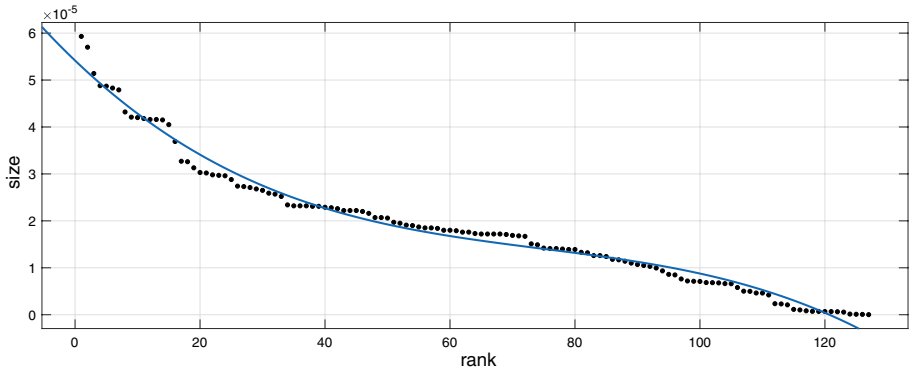
Let us explore now the subsamples.

When dealing with the 10% of companies with highest clustering coefficient, we observe that in both cases of Mafia and No Mafia companies we have a decrease of the distance between the values of the entropy and the maximum one, so that the distribution of this subsample is quite different from the uniform one. Such a behavior is much more evident in the No Mafia case, with a value of entropy close to zero. This outcome is a further confirmation that Mafia companies are more scattered than No Mafia ones in terms of the creation of risk-based community structures, even in the case of highly community levels. Moreover, this finding is also confirmed when taking the 20% and 50% of companies with highest clustering coefficient. In particular, the values of the ratio between the entropy of the overall sample and the maximum possible entropy are quite close to those of the 50% highest subsample case. This suggests that the behavior of the overall sample in terms of shape of the distribution of the clustering coefficients replicates the one of its half-part subsample with the highest values of the clustering coefficients.

The subsamples with lower levels of clustering coefficients are quite similar in the Mafia and No Mafia cases. The entropy is quite far from its maximum possible levels in all the cases of 10%, 20% and 50% subsamples, and such a distance decreases with respect to the size of the subsample. It is worth to point out that also in these cases Mafia companies are generally more scattered than the No Mafia ones. In the limit situation of the 10%



**Fig. 1** Best fit third-degree polynomial curve as in (7) of the clustering coefficients of the Mafia companies. To have a better visual inspection of the goodness of fit, scatter plot and calibrated curve are juxtaposed. On the *x*-axis we have the ranks, while on the *y*-axis the size in terms of clustering coefficient. The best fit parameters and the calibrated parameters can be found in Table 8



**Fig. 2** Best fit third-degree polynomial curve as in (7) of the clustering coefficients of the No Mafia companies. For details, see the caption of Fig. 1. Best fit and calibrated parameters are in Table 9

**Table 8** Calibrated parameters of the rank-size function in (7) for the case of Mafia companies

a	$-1.251e-09$ ( $-1.521e-09, -9.816e-10$ )
b	$2.419e-07$ ( $2.017e-07, 2.821e-07$ )
c	$-1.574e-05$ ( $-1.744e-05, -1.404e-05$ )
d	$0.000377$ ( $0.0003576, 0.0003963$ )
SSE	$4.934e-08$
R-square	0.9422
RMSE	$2.303e-05$

In the brackets, the confidence bounds at 95%. The goodness of fit parameters SSE, R-square and RMSE are also presented

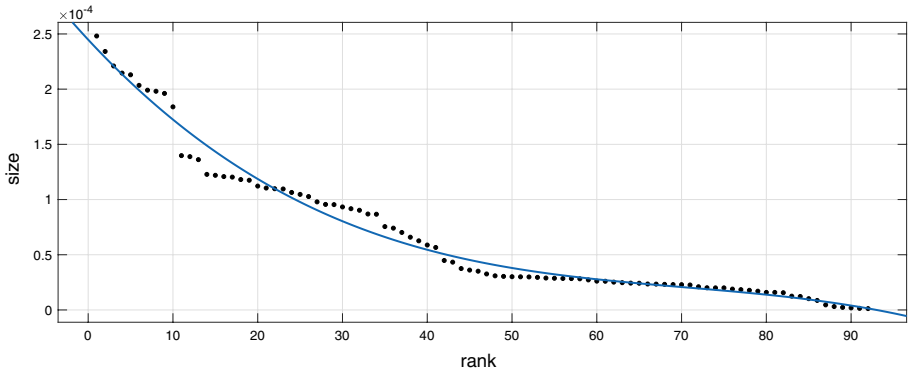
of companies with the lowest levels of clustering coefficients, the values of entropies are particularly close to zero, so that the companies with the weakest community structures exhibit no substantial differences in their community structure levels.



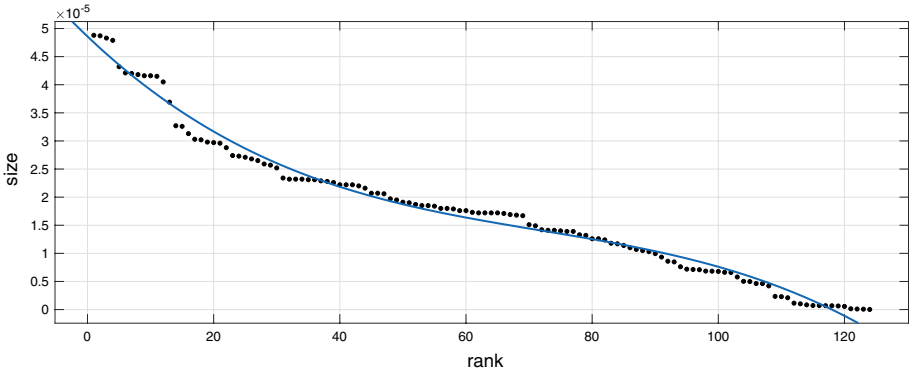
**Table 9** Calibrated parameters of the rank-size function in (7) for the case of No Mafia companies

a	$-6.564e-11$ ( $-7.345e-11$ , $-5.783e-11$ )
b	$1.474e-08$ ( $1.322e-08$ , $1.626e-08$ )
c	$-1.271e-06$ ( $-1.355e-06$ , $-1.187e-06$ )
d	$5.416e-05$ ( $5.291e-05$ , $5.54e-05$ )
SSE	$3.639e-10$
R-square	0.9841
RMSE	$1.72e-06$

For the details on the content of the Table, please read the caption of Table 8



**Fig. 3** Best fit third-degree polynomial curve as in (7) of the clustering coefficients of the Mafia companies when one king and four vice-roys are removed from the sample. The related best fit and calibrated parameters are presented in Table 10



**Fig. 4** Best fit third-degree polynomial curve as in (7) of the clustering coefficients of the No Mafia companies when one king and two vice-roys are removed from the sample. Best fit and calibrated parameters are given in Table 11

**Table 10** Calibrated parameters, confidence bounds at 95% and goodness of fit parameters of the rank-size function in (7) for the case of Mafia companies and when one king and four vice-roys are removed from the sample

a	$-5.126e-10$ ( $-6.391e-10$ , $-3.86e-10$ )
b	$1.083e-07$ ( $9.036e-08$ , $1.262e-07$ )
c	$-8.268e-06$ ( $-8.987e-06$ , $-7.549e-06$ )
d	$0.0002448$ ( $0.0002371$ , $0.0002526$ )
SSE	$7.1e-09$
R-square	$0.9807$
RMSE	$8.982e-06$

For the details on the content of the Table, please read the caption of Table 8

**Table 11** Calibrated parameters, confidence bounds at 95% and goodness of fit parameters of the rank-size function in (7) for the case of No Mafia companies and when one king and two vice-roys are removed from the sample

a	$-5.667e-11$ ( $-6.405e-11$ , $-4.93e-11$ )
b	$1.225e-08$ ( $1.084e-08$ , $1.365e-08$ )
c	$-1.068e-06$ ( $-1.143e-06$ , $-9.921e-07$ )
d	$4.858e-05$ ( $4.749e-05$ , $4.968e-05$ )
SSE	$2.676e-10$
R-square	$0.9857$
RMSE	$1.493e-06$

For the details on the content of the Table, please read the caption of Table 8

## 5.2 Rank-size analysis

The best fit curves procedures here proposed have been implemented through the Matlab software, by using the cftool toolkit and—according to equation (7)—by selecting a polynomial law of degree three.

We present in Figs. 1 and 2 the best fit curves juxtaposed to the scatter plot of the real data, for a better visualization of the final outcomes, in the cases of Mafia and No Mafia companies. Tables 8 and 9 contain the calibrated parameters, along with the goodness of fit  $R^2$ , in both of the considered cases.

In both of cases of Mafia and No Mafia companies we observe rather satisfactory goodness of fit parameters (see Tables 8 and 9, the last three rows). However, the visual inspection of how the calibrated curves offer a good representation of the scatter plots suggests the presence of five outliers at high ranks (one king and four vice-roys) for the Mafia companies and three outliers (one king and two vice-roys) for the No Mafia ones. The removal of the king and vice-roys from the samples leads to more satisfactory visual appeal of the best fit curve, and also an improvement of the goodness of fit parameters—being such an improvement more evident for the Mafia companies. See Figs. 3 and 4 and the corresponding Tables 10 and 11 for the Mafia companies and No Mafia ones cases, respectively.

As a general comment, we can argue that the considered samples of companies can be well-represented through unified curves when dealing with their community structures in terms of closeness of the risk profiles. This outcome can be associated to the possibility of moving from a scatter plot of observations at microscopic level to a best fitting curve at a more systemic level. Substantially, companies are able to generate a system, so that one can theoretically identify the value of the clustering coefficients of a hypothetical company

**Table 12** Value of the calibrated  $-b/3a$  in the four cases of Mafia and No Mafia companies, for the full sample and when king and vice-roys are removed

	Mafia	No Mafia	Mafia (no king and vice-roys)	No Mafia (no king and vice-roys)
$-b/3a$	64.46	74.85	70.43	72.05
Sample dimension	97	127	92	124
$(-b/3a) / \text{Sample dimension}$	0.66	0.59	0.77	0.58

The ratio between  $-b/3a$  and the dimension of the considered sample is also included, for an easy interpretation of these results

which is inserted in the sample with a specific rank. The presence of outliers at high ranks lets this procedure be less powerful, in that king and vice-roys effects are able to drive the best fit procedures towards curves with higher average distances from the observed samples. In this respect, the removal of the outliers is able to offer more convincing results under a statistical point of view.

Further insights on the interpretation of the results can be derived by understanding the meaning of the parameters  $a$ ,  $b$ ,  $c$ ,  $d$  in formula (7). In this respect, the value of  $d$  represents the intercept of the curve with the  $y$ -axis. Therefore, the size of the clustering coefficient of the network at the highest rank – more specifically, at the hypothetical rank  $r = 0$ —increases as the value of  $d$  increases. The value of  $c$  describes the slope of the straight line which is tangent at the curve at the highest rank  $r = 0$ . Hence, such a value gives insights on the relationships between companies at high consecutive ranks. By construction we have  $c < 0$ , and the distance between the sizes at ranks 1 and 2 increases as the absolute value of  $c$  increases. For the parameters  $a$  and  $b$ , we find convenient to discuss the ratio  $-b/3a$ . Indeed, we have a change of shape of the curve from convexity to concavity at  $r = -b/3a$ . A high value of  $-b/3a$  is associated to a low rank when passing from large differences between consecutively ranked companies in terms of community structures associated to risk profiles—convex behavior of the curve—to small differences between them. More insights can be derived by analyzing the ratio between the change point of the shape of the curve and the cardinality of the considered sample. Please, refer to Table 12 for the analysis of this aspect.

Tables 8, 9, 10 and 11 offer a clear view of the results best fit exercises, which can be interpreted in the context of Mafia and No Mafia companies.

As intuition suggests, the value of  $d$  is higher when king and vice-roys are not removed, in both of cases of Mafia and No Mafia companies. However, the ones of Mafia companies are much higher than those of the No Mafia ones. Such an outcome points to attention the much higher community structures in terms of closeness of risk profiles in the Mafia case than in the No Mafia one, when highest ranks are considered. The absolute value of  $c$  is remarkably higher for the case of Mafia rather than in the No Mafia one—more than ten times more in presence of king and vice-roys and about eight times more otherwise. This outcome suggests that the highest ranked companies have a much higher distance in the case of Mafia than in the No Mafia one. A joint analysis of  $c$  and  $d$  offers a view of the Mafia companies much more scattered than the No Mafia ones in correspondence of the nodes with the highest levels of community structures. The ratio between  $-b/3a$  and the dimension of the considered sample gives

that companies tend to become more heterogeneous for lower values of the ranks—with a concave shape of the curve—in the Mafia case than in the No Mafia one. Moreover, there is not a substantial difference in the No Mafia case when king and vice-roys are removed or not, while Mafia companies tend to be more heterogeneous at very low levels of the ranks when king and vice-roys are removed. This finding is a further support of the more prominent differences—to be intended in terms of clustering associated to the common risk profiles—between Mafia companies than No Mafia ones, mainly when high levels of community structures are explored.

The results suggest that, under some conditions, Mafia companies hold a risk profile remarkably higher than No Mafia ones, which undoubtedly can provide a significant alert for the players directly involved in the operating activities of the companies, as well as for the whole economic and social communities that Mafia companies belong to. Anyway, the level of heterogeneity among Mafia companies may be interpreted as a further evidence for the existence of companies of different kind, holding different characteristics; as a consequence, this produces a remarkable variability in their risk profiles. Conversely, all No Mafia companies are reasonably expected to be managed in the pursue of improvements in their economic and financial performance, which may tend to produce a larger extent of homogeneity. A deeper analysis—with refined clustering procedures – could be an interesting development of the rank-size analysis for providing additional insights about the multiple types of Mafia companies.

## 6 Conclusions

This paper provides an evaluation of the companies publicly registered as Mafia ones, with a specific focus on the distribution of their financial risk profiles. More than this, it discusses the communities that companies implicitly form when they have similar financial risk profiles of high entity.

At this aim, we present the statistical comparison between Mafia companies with No Mafia ones in a complex network framework. The weights of the arcs of the networks are defined by introducing an overall risk indicator, which synthesizes the financial risk ratios of the individual companies. Communities are suitably measured through the computation of the clustering coefficients at nodes level. The analysis is carried out under two different perspectives: by one side, we implement a rank-size best fit procedure; by the other side, we provide an entropy-based discussion of the distribution of the clustering coefficients.

Empirical experiments are grounded on a high level dataset of Italian companies, which can be effectively clustered into the Mafia and No Mafia groups.

Results show that Mafia companies have higher financial risk profiles and are more heterogeneous than the No Mafia ones, mainly at high risk level. Therefore, the research hypothesis **H1** is confirmed to be satisfied. We can say more than this. Indeed, the rank-size analysis highlights that the observed set of companies form two universal systems, whose usefulness in describing the risk-based properties of the overall environment is clear. In this respect, the detection of remarkable outliers at high levels of nodes clustering coefficients—more evident in the network of Mafia companies—points to the presence of distortion factors for the best fit procedure.

The paper provides several contributes to the literature as well as some practical implications.

On the practical side, the proposed methodology can be used by banks and financial institutions as well as by regional or local authorities, in order to monitor companies' risk profiles and get relevant information—not only under the perspective of detection of Mafia companies. Indeed, the methodology could be adopted, as an example, by clustering companies by industry and define suitable supporting initiatives. In the field of study of criminal organizations, the results provide further evidences to extant literature (Ravenda et al. 2015a, b; Fabrizi et al. 2017) about the commonalities that may differentiate Mafia and No Mafia Companies under a risk perspective. A further development could take into consideration a refinement of Mafia company clustering, based on the different types of Mafia companies, which could provide a more comprehensive description of the differences in the associated risk profiles. By the way, analyzing in depth the different subcategories might provide room for the construction of more accurate risk profile indexes, including profitability measures as well as other kinds of financial ratios where suitable. In this respect, we also point out that the cardinality of the considered sample is not large enough to allow exploration at a regional level—which is undoubtedly interesting. Specifically, a regional analysis of the Italian companies requires more data and/or the employment of different investigation methodologies.

Under a purely methodological perspective, the paper aims to widen the adoption of a complex network approach to new real world phenomenon conditioned by relevant ambiguity and complexity, then extending previous contributions of Cinelli et al. (2016, 2017, 2020); Cerqueti et al. (2018); D'Arcangelis et al. (2020), only to name a few. In this respect, this paper can be extended in several directions. Indeed, starting from the proposed statistical-methodological framework, we are able to detect the presence of more prominent companies and derive, accordingly, the topological structure of the networks. Moreover, the exploration of the assortativity property of the networks (Arcagni et al. 2017) can provide further details on the way in which companies are linked, on the basis of their financial risk profile. We leave all these challenging explorations to future research.

In a more general framework, we can identify other patterns of research directions. On one side, one can explore the performance-related to takeovers related mafia-related companies in all the available cases of Mafia/No Mafia company takeover Mafia/No Mafia company. On the other side, it might be interesting to compare Mafia companies to other Organized Crime types in different regional realities, like the Keiretsu companies in Japan. In doing so, one needs to perform a deep preliminary exploration of the related business activities' similarities.

**Funding** Open access funding provided by Università di Pisa within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahmed K, Courtis JK (1999) Associations between corporate characteristics and disclosure levels in annual reports: a meta-analysis. *Br Account Rev* 31:35–61
- Allum F, Marinaro IC, Sciarrone R (eds) (2019) *Italian mafias today: territory, business and politics*. Edward Elgar Publishing
- Arcagni A, Grassi R, Stefani S, Torriero A (2017) Higher order assortativity in complex networks. *Eur J Oper Res* 262:708–719
- Arlacchi P (2007) *La mafia imprenditrice. Dalla Calabria al centro dell'inferno*. Milano, Il sagggiatore
- Arlacchi P, Ryle M (1986) *Mafia business: the mafia ethic and the spirit of capitalism*, vol 3, p 4
- Ausloos M (2013) A scientometrics law about co-authors and their ranking: the co-author core. *Scientometrics* 95:895–909
- Ausloos M, Cerqueti R (2016) A universal rank-size law. *PLoS One* 11(11):e0166011
- Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A (2004) The architecture of complex weighted networks. *Proc Nat Acad Sci* 101(11):3747–3752
- Bartolacci F, Castellano NG, Cerqueti R (2015) The impact of innovation on companies' performance: An entropy-based analysis of the STAR market segment of the Italian Stock Exchange. *Technol Anal Strateg Manag* 27(1):102–123
- Bartolacci F, Cerqueti R, Paolini A, Soverchia M (2019) An economic efficiency indicator for assessing income opportunities in sustainable waste management. *Environ Impact Assess Rev* 78:106279
- Bartram SM, Wang YH (2005) Another look at the relationship between cross-market correlation and volatility. *Finan Res Lett* 2(2):75–88
- Baucus MS, Near JP (1991) Can illegal corporate behavior be predicted? An event history analysis. *Acad Manag J* 34:9–36
- Bera AK, Park SY (2008) Optimal portfolio diversification using the maximum entropy principle. *Econom Rev* 27(4–6):484–512
- Berlusconi G (2016) Organised criminals and the legal economy. In Savona EU, Riccardi M, Berlusconi G (eds) *Organised crime in European businesses*, Chapter 1. Routledge, London, pp 4–15
- Briant A, Combes PP, Lafourcade M (2010) Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *J Urban Econ* 67(3):287–302
- Calderoni F (2011) Where is the mafia in Italy? Measuring the presence of the mafia across Italian provinces. *Glob Crime* 12(1):41–69
- Cao Z, Leng F, Feroz EH, Davalos SV (2015) Corporate governance and default risk of firms cited in the SEC's Accounting and Auditing Enforcement Releases. *Rev Quant Financ Acc* 44:113–138
- Castellano N, Del Gobbo R, Corsi K (2019) Analysis of disclosure determinants: a local-relation approach. *Medit Account Res* 27(3):399–415
- Cerqueti R, Ausloos M (2015a) Cross ranking of cities and regions: population versus income. *J Stat Mech: Theory Exp* 2015(7):P07002
- Cerqueti R, Ausloos M (2015b) Evidence of economic regularities and disparities of Italian regions from aggregated tax income size data. *Physica A* 421:187–207
- Cerqueti R, Clemente GP, Grassi R (2020) Stratified communities in complex business networks. *J Bus Res*. <https://doi.org/10.1016/j.jbusres.2020.04.005>
- Cerqueti R, Ferraro G, Iovanella A (2018) A new measure for community structures through indirect social connections. *Expert Syst Appl* 114:196–209
- Chao A, Jost L, Hsieh TC, Ma KH, Sherwin WB, Rollins LA (2015) Expected Shannon entropy and Shannon differentiation between subpopulations for neutral genes under the finite island model. *PLoS One* 10(6):e0125471
- Cinelli M, Ferraro G, Iovanella A (2016) Some insights into the relevance of nodes' characteristics in complex network structures. In Zylka M, Fuehres H, Fronzetti Colladon A, Gloor P (eds) *Designing Networks for Innovation and Improvisation*. Springer, Cham, pp 77–84
- Cinelli M, Ferraro G, Iovanella A (2017) Structural bounds on the dyadic effect. *J Complex Netw* 5(5):694–711
- Cinelli M, Peel L, Iovanella A, Delvenne JC (2020) Network constraints on the mixing patterns of binary node metadata. *Phys Rev E* 102(6):062310
- Cinelli M (2020) Ambiguity of network outcomes. *J Bus Res*. <https://doi.org/10.1016/j.jbusres.2020.02.031>
- Clemente GP, Grassi R (2018) Directed clustering in weighted networks: a new perspective. *Chaos Solitons Fract* 107:26–38
- Cooke TE (1989) Disclosure in the corporate annual reports of Swedish companies. *Account Bus Res* 19:113–124

- Cressey DR (1976) Restraint of trade, recidivism and delinquent neighborhoods. In: Short JF (ed) *Delinquency, crime and society*. University of Chicago Press, Chicago, pp 209–238
- D’Arcangelis AM, Levantesi S, Rotundo G (2019) A complex networks approach to pension funds. <https://doi.org/10.1016/j.jbusres.2019.10.071>
- D’Arcangelis AM, Rotundo G (2019) Herding in mutual funds: a complex network approach. *J Bus Res*. <https://doi.org/10.1016/j.jbusres.2019.11.016>
- Daboub AJ, Rasheed AM, Priem RL, Gray D (1995) Top management team characteristics and corporate illegal activity. *Acad Manag Rev* 20(1):138–170
- Di Bono L, Cincimino S, Riccardi M, Berlusconi G (2015) Management strategies of infiltrated businesses. In: Savona EU, Berlusconi G (eds) *Organized crime infiltration of legitimate businesses in Europe: A pilot project in five European countries*. Transcrime - Università degli Studi di Trento, Trento, pp 102–112
- Dupla V, Very P, Monnet B (2012) Identification and economic analysis of governance mechanisms in legally registered mafia firms. *Mngement* 15:273–282
- Elliott M, Golub B, Jackson MO (2014) Financial networks and contagion. *Am Econ Rev* 104(10):3115–3153
- Esposito P, Brusoni M, Borgonovi E (2019) Public integrity in Italy. *Mastering Anti-Corruption. The Practitioners’ View*, p 75
- Fabrizi M, Malaspina P, Parbonetti A (2017) Caratteristiche e modalità di gestione delle aziende criminali. *Rivista di Studi e Ricerche sulla criminalità organizzata* 3(1):47–66
- Fagiolo G (2007) Clustering in complex directed networks. *Phys Rev E* 76(2):1
- Fedajev A, Stanujkic D, Karabašević D, Brauers WK, Zavadskas EK (2020) Assessment of Progress towards “Europe 2020” strategy targets by Using the MULTIMOORA Method and the Shannon Entropy Index. *J Clean Prod* 244:118895
- Firth M, Smith A (1995) Auditor Quality, Corporate Risk, and the Valuation of New Issues. *Rev Quant Financ Acc* 5:241–251
- Forbes KJ, Rigobon R (2002) No contagion, only interdependence: measuring stock market comovements. *J Finan* 57(5):2223–2261
- Gabaix X (1999a) Zipf’s law for cities: an explanation. *Q J Econ* 114(3):739–767
- Gabaix X (1999b) Zipf’s Law and the Growth of Cities. *Am Econ Rev* 89(2):129–132
- Gabaix X (2009) Power laws in economics and finance. *Ann Rev Econ* 1(1):255–294
- Grassi R, Calderoni F, Bianchi M, Torriero A (2019) Betweenness to assess leaders in criminal networks: new evidence using the dual projection approach. *Soc Networks* 56:23–32
- Karagiannis R, Karagiannis G (2020) Constructing composite indicators with Shannon entropy: the case of human development index. *Socio-Econ Plan Sci* 70:100701
- Koutmos D, Bozos K, Dionysiou D, Lambertides N (2018) The timing of new corporate debt issues and the risk-return tradeoff. *Rev Quant Financ Acc* 50:943–978
- La Rosa F, Paternostro S (2015) L’amministrazione delle aziende confiscate alla criminalità organizzata. La proposta di un framework di studio e un’analisi esplorativa dei decreti di destinazione. *Rivista Piccola Impresa/Small Business* 1:54–84
- La Rosa F, Paternostro S, Picciotto L (2018) Exploring the determinants of anti-mafia entrepreneurial behaviour: an empirical study on southern Italian SMEs. *Entrepren Region Dev* 30(1–2):81–117
- Laherrere J, Sornette D (1998) Stretched exponential distributions in nature and economy fat tails with characteristic scales. *Eur Phys J B* 2(4):525–539
- Mastrobuoni G, Patacchini E (2012) Organized crime networks: an application of network analysis techniques to the American mafia. *Rev Network Econ* 11(3):1
- Mirenda L, Mocetti S, Rizzica L (2019) The real effects of ‘ndrangheta: firm-level evidence. *Bank of Italy Working Papers*, No. 1235
- Mollgaard A, Zettler I, Dammeyer J, Jensen MH, Lehmann S, Mathiesen J (2016) Measure of node similarity in multilayer networks. *PLoS One* 11(6):e0157436
- Onnela J, Saramaki J, Kertesz J, Kaski K (2005) Intensity and coherence of motifs in weighted complex networks. *Phys Rev E* 71(6):1
- Ozgul F (2016) Analysis of topologies and key players in terrorist networks. *Socio-Econ Plan Sci* 56:40–54
- Paoli L (2004) Italian organised crime: Mafia associations and criminal enterprises. *Global Crime* 6(1):19–31
- Pinotti P (2015) The economic costs of organised crime: Evidence from Southern Italy. *Econ J* 125(586):F203–F232
- Pola G (2016) On entropy and portfolio diversification. *J Asset Manag* 17(4):218–228
- Ramchand L, Susmel R (1998) Volatility and cross correlation across major stock markets. *J Emp Finan* 5(4):397–416

- Ravenda D, Argiles-Bosch JM, Valencia-Silva MM (2015a) Labor Tax Avoidance and Its Determinants: The Case of Mafia Firms in Italy. *J Bus Ethics* 132(1):41–62
- Ravenda D, Argiles-Bosch JM, Valencia-Silva MM (2015b) Detection model of legally registered mafia firms in Italy. *Eur Manag Rev* 12(1):23–39
- Rosser JB (2016) Entropy and econophysics. *Eur Phys J Spec Top* 225(17–18):3091–3104
- Savona E, Berlusconi G (2015) Organized crime infiltration of legitimate businesses in Europe: a pilot project in five European countries. Final report of project ariel—assessing the risk of the infiltration of organized crime in EU MSS legitimate economies: a pilot project in 5 EU countries
- Savona EU, Calderoni F (Eds) (2016) Criminal markets and mafia proceeds. Routledge
- Savona E, Riccardi M (2018) Mapping the Risk of Serious and Organised Crime Infiltration in European Businesses-Final Report of the MORE Project. Transcrime-Università Cattolica del Sacro Cuore, Milan
- Sciarrone R (2009) Mafie vecchie mafie nuove. Radicamento ed espansione. Donzelli, Roma
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423
- Simpson SS (1986) The decomposition of antitrust: Testing a multi-level, longitudinal model of profit-squeeze. *Am Sociol Rev* 51:859–875
- Transcrime (2013) Progetto PON (Programma operativo nazionale) Sicurezza 2007-2013: Gli investimenti delle mafie. Rapporto Linea 1. Ministero dell'Interno, Milano. <http://www.investimentioc.it/>
- Villani S, Mosca M, Castiello M (2019) A virtuous combination of structural and skill analysis to defeat organized crime. *Socio-Econ Plan Sci* 65(C):51–65
- Watts DJ, Strogatz SH (1998) Collective dynamics of small-world networks. *Nature* 393(6684):440–442
- Yang J (2018) Information theoretic approaches in economics. *J Econ Surv* 32(3):940–960

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.