Dipartimento di Economia, Ingegneria, Società e Impresa

Corso di laurea in

Amministrazione, Finanza e Controllo

”Business Networks: an Empirical Analysis”

Insegnamento

Economia Industriale

RELATORE
Prof.re Luca Correani

CANDIDATO
Valentina Ceccarelli 52977

Anno Accademico 2021/2022

Alla mia famiglia

Acknowledgements

Giunta a conclusione del mio percorso universatario, colgo l'occasione per ringraziare tutti coloro che durante questi anni hanno sempre avuto fiducia in me, sostenendomi giorno dopo giorno.
Vorrei anche ringraziare, innanzitutto, me stessa per non aver mai mollato, aver creduto sempre in me e per essere arrivata fino a questo traguardo. Questi anni, non solo mi hanno formato come studentessa, ma anche come persona, rendendomi più matura e consapevole.
Un ringraziamento particolare va al mio relatore, il professor Luca Correani, che mi ha dato questa l'oppurtunità e mi ha seguito durante il mio ultimo lavoro da studentessa universataria, stimolando un vivo interesse per la materia. Vorrei, inoltre, ringraziare tutti i miei docenti che durante i corsi hanno condiviso la loro esperienza con passione e dedizione, in particolar modo, ricordo con molto piacere il corso di economia e globalizzazione dello sviluppo svolto dal professor Giulio Guarini che mi ha dato l'opportunità di riflettere su quanto la formazione possa essere strumento essenziale per migliorare la nostra persona e le nostre aspettative future.
Vorrei ringraziare anche tutti i miei amici che hanno condiviso con me momenti di spensieratezza e allegria, in particolare le mie migliori amiche, Martina e Marika, che mi sono sempre state vicino sopratutto nei momenti di bisogno.
Vorrei ringraziare la mia famiglia composta dai miei genitori e da mio fratello. Ad Andrea, vorrei dire grazie per tutto il bene che mi dimostra ogni giorno e per esserci sempre. Sono infinitamnete grata ai miei genitori: mi hanno dato, infatti, l'opportunità di proseguire gli studi e hanno facilitato, sostenendomi e incoraggiandomi, il mio percorso universitario.
Dal più profondo, voglio ringraziare Gianmarco, mio fidanzato e giovane brillante ricercatore, che durante questo percorso ha saputo consigliarmi, incoraggiarmi, starmi accanto e riprendermi al momento opportuno. Fin da subito, ha dismostrato, infatti, un vivo e sincero interesse per i miei studi e per le mie ricerche, risvegliando in me curiosità e sete di conoscenza. Ad oggi, non riesco ad immaginare la mia vita senza l'amore e i bei momenti che questo giovante e intelligente Uomo mi dona ogni giorno, spero di condividere con lui molti altri bei momenti come questo!
Voglio concludere questi ringraziamenti citando Antoine De Saint-Eupèry che ne "Il piccolo principe" sciveva: "Ecco il mio segreto. É molto semplice: non si vede che con il cuore. L'essenziale 'e invisibile agli occhi". *Ad maiora!*

Valentina

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Network companies, or business networks, are a form of aggregation of different companies that aim to cooperate to acquire new managerial and technical skills and improve effectiveness to carry out several activities, which alone would not be able to accomplish. In the literature, there are different definitions of network companies; as an example in [Garofalo and Pugliesi, 2014], they are intensive long-term collaborations that aim to strengthen competitiveness and innovative capacity sharing the results, and also maintaining its legal and economic autonomy. Companies may cooperate as only one large corporation with high skills, innovations, and risk-taking abilities. Indeed, belonging to a network allows them to access new markets, improve their competitive advantages, and access economies of scale and variety typical of big companies.

In Italy, it is possible to distinguish two forms of collaboration: one is formal, and the other is informal. Formal collaborations, also known as "Reti soggetto", distinguish from informal partnerships because the first has legal subjectivity and the second has not got. "Reti soggetto" have an own autonomy. Formal collaboration is defined as a contract between companies, mandatorily endowed with an asset fund and a common organ that acquires independent legal personality with registration in the ordinary section of the Register of Companies. A "Rete contratto" is more like an actual contract between enterprises without acquiring legal subjectivity. Informal collaboration is a contract that enables partnerships based on a common program and strategic competitiveness goals; that is, collaborative and sharing relationships are formalized to define the commitment, investment, and type of bonding to be adopted [Rizzo, 2020].

Legally, they are born and evolved with the following legislation: Decree-Law No. 5

of 2009, converted into Law No. 33 of 2009 and amended by Law No. 99 of 2009, Decree-Law No. 78 of 2010 ratified in Law No. 122 of 2010, Decree-Law No. 83 of 2012 converted into Law No. 134 of 2012, Decree-Law No. 179 of 2012 ratified in Law No. 221 of 2012 and, Law No. 154 of 2016. These are the most important laws related to network contracts.

The purpose of this thesis is to study informal business networks in Italy. In Italy, economic growth is not encouraged by the production system that is mainly small and medium-sized enterprise-based. Small-Medium Enterprises (SMEs) do not facilitate economic growth and are a huge limitation of the Italian production system. The Italian growth rate's trend is decreasing, as shown in the World Bank Database [World Bank Group, 2022]. Firms' dimension is one of the factors limiting economic growth.

The Total Factor Productivity (TFP), as a measure of technological advancement, is decreasing [University of Groningen and University of California, Davis, 2022b]; one of the reasons is that in Italy, firms are mainly small. One of the solutions to this problem is for companies to decide to collaborate and cooperate. The deficit of innovation is a weakness of Italian industry: it is highly correlated with the firm's dimension. Innovation raises the productivity with efficient infrastructure, job markets, innovative organizations, and a constant collaboration from the private and public sectors [Garofalo and Pugliesi, 2014].

The dimension is a significant restriction for competitiveness and internationalization, as the market is globalized and there are no physical boundaries. A small firm does not often have the proper resources to compete with the market, but it has rapidity in decisions, flexibility and the ability of customized productions [Aureli and Del Baldo, 2014]. All kinds of business aggregations are significant because they enable them to overcome all the limits typical of SMEs. In particular, networks are relevant because they aim to strengthen competitiveness, especially in foreign markets, to improve economic performance and to achieve the economies of scale and "system" typical of medium and large enterprises without giving up the advantages and flexibility of small-size [Zazzaro, 2010]. Internationalization, competitiveness, innovation and dimension are weak points of the Italian production system.

R&D is a typical subject of economic growth literature. In this theoretical framework, a [Di Dio et al., 2022] model has been applied. From a micro-economic perspective, the model wants to investigate the relationship between the nature of goods produced by firms operating in the same R&D network and their strategic location. Especially, the paper's main point is to understand if there is a matching between the average distance of a business network and the type of good produced by the same network. A spatial Cournòt model was implemented taking into account strategic location, R&D investment, and differentiated production. This thesis starts from the work by [Di Dio et al., 2022], which is a theoretical study that analyses only networks with two companies. It extends the treatment of informal networks in all the country. An empirical analysis was conducted on the dataset by the Chamber of Commerce [Camere di Commercio d'Italia, 2022]. The main contribution of this thesis consists of an empirical analysis of business networks performed through a cross-section regression. For each network, the most important factors are the average distance, the kind of produced good, the production sector, the geographic region, and the dimension of the network.

The next chapter presents a general introduction to business networks in Italy. It analyses reports and related literature. In particular, our focus will be on several variables. The average number of companies indicates the network's dimension, and represents a development indicator of the network's complexity. As the phenomenon has developed over time, considering the date of establishment of the network contracts is crucial to understanding the depth of the phenomenon. Other important variables are the production sector, following the Ateco 2 Digit convention by Istat [Istat, 2022b], the macro-area production and the region of business networks. These three latter parameters are computed not only for the networks but also for each company, to see which kind of firms decides to agree to the network's contracts. Other descriptive statistics may be included in the section.
In the third chapter, the regression model will be applied to the [Camere di Commercio d'Italia, 2022] dataset considering all the business networks with two firms and then, on the entire dataset to evaluate the theoretical model from [Di Dio et al., 2022]. In the regression model, we try to understand if the dependent variable, which is the average distance of each network, may affect the independent ones. They are: the network geographic region, the type of good or service produced, the dimension of the network stands for the number of companies in each network, and the productivity sector. The target is try to answer to the following question: Can the physical distance between companies affect product differentiation in business networks? It should be noted that the cross-section regression model does not consider the time dimension.

Chapter 4 reports the original contribution of this thesis, the conclusions and possible future works.

Finally, the A and B sections report the appendices; in the A appendix, there is the Ateco code classification, and in the B appendix the regression results of the empirical model.

Chapter 2

Business Networks in Italy

The second chapter aims to provide a general background of Italian business networks. The overview starts from a macroeconomic to a microeconomic perspective.
The first paragraph regards globalization and international trade, the second describes foreign direct investment and capital flows to Italy, the third analysis TFP, and then it wants to present the strengths and the weakness of the Italian production system. From a microeconomic perspective, the chapter proceeds with the differences between industrial districts and business networks, highlighting the importance of business networks and the member's characteristics of the impact of the informal partnerships. In conclusion, you can find out about a networks description and firms by date, average number, region, macro-area of activity and business sector.

2.1 Globalization and International Trade

"Globalization is a process that encompasses the causes, course, and consequences of transnational and trans-cultural integration of human and non-human activities." [Al-Rodhan and Stoudmann, 2006]. As a result, globalization affects our lives and our mindset. There are no more physical boundaries. Indeed, differences between people are thinned out.
In the last fifty years, globalization has caused many changes, especially in the production sector, as it drives economic growth. Indeed, economies of scale are a requirement for companies to access the globalized international trade market. There are two kinds

of economies of scale: internal and external. Single firms produce the internal ones; instead, the external ones occur at the sector level.

Marshall was the first who studied the industrial district [Marshall, 2009]. He sustains that a cluster of firms may be more efficient than a single firm because a cluster can attract specialized suppliers, generate a pool of workers with proper qualifications, and promote knowledge spillover [Krugman and Obstfeld, 2009]. International trade is more efficient with collaboration and cooperation between businesses, but some other factors or institutions are able to affect international trade such as a country's government. [Milner, 1999] tries to answer how an institution may impact international trade. As an example, the pressure group model has attempted to delineate more specifically the groups who should favor and oppose protection, and the conditions under which they may be most influential [Milner, 1999].

It is evident that in this kind of countries, factors of productivity, such as labour or capital, are scarce. In an economy, where export-oriented industries and multinationals tend to favor freer trade and to be associated with less protection, there is an increment of productivity. In Less Development Countries (LDCs), an opening to international trade prevents economic crisis. In LDcs, trade policy is affected by domestic factors and external institutions such as World Trade Organization (WTO), International Monetary Fund (IMF) and World Bank [Milner, 1999]. These organizations have added leverage to arguments for trade liberalization, but they only give general guidelines. Globalization often makes possible governments cut spending on social programs and reduces taxes on capital [Milner, 1999]. Furthermore, a country with high international trade develops a high rate of international relations, promoting peace and stability.

Developed Countries (DCs) should reduce the barrier to entry in their market for LDCs countries as international trade establishes a virtuous cycle of growing trade creating more groups in favor of trade liberalization, which in turn created more impetus for greater liberalization, and more trade seems to be a key factor [Milner, 1999].

The export firm is more productive, skill-and capital-intensive [Bernard et al., 2007]. Exporters have a productivity advantage before starting to export suggests self-selection, export firms are more productive and resilient. This microeconomic evidence influences macroeconomic output.

The role of distance is in dampening aggregate trade flows [Bernard et al., 2007]. A combination of economies of scale and consumer's preferences for variety leads otherwise identical firms to "specialize" in distinct horizontal varieties, spurring two-way or “intra-industry” trade between countries [Bernard et al., 2007].

Exporting produces productivity and companies ”learn by exporting”, which means they increment their value. The welfare gains from trade (a.k.a. comparative advantage theory) and the combination of economies of scale with expanding product varieties available to consumers [Milner, 1999]. International trade is very concentrated between firms because they tend to trade with similar countries and firms. It contributes firstly to increment inequality in international trade, and then in corresponding countries.

2.2 Foreign Direct Investment and Capital Flows

Foreign Direct Investments (FDIs) are international investments within the balance of payment accounts. Essentially, a resident company in one economy seeks to obtain a lasting interest in an enterprise resident in another economy [Eurostat, 2018]. A lasting interest implies the existence of a long-term relationship between the direct investor and the enterprise, and an investor's significant influence on the enterprise management. An enterprise's direct investment is one in which a direct investor owns 10% or more of the ordinary shares or voting rights (for an incorporated enterprise) or the equivalent (for an unincorporated enterprise) [Eurostat, 2018].

For Italy, FDIs are important, in particular, [Bronzini, 2007] finds empirical evidence on the FDIs impact of localization economies. Foreign investors are interested in merging with or acquiring big companies to increment market shares in the host country [Bronzini, 2007], usually less available in small firms; at the same way, small firms may be more attractive because they are dynamic, thanks to the geographical location, the size of Italian firms and the presence of industrial districts. According to Marshall, the spatial concentration of firms producing similar good provides positive externalities, such as knowledge spillover and labor pooling. Sharing of ideas and knowledge of new technologies apparently spreads more rapidly across firms that are concentrated in specialized areas, due to informal contacts, networks or the mobility of workers across firms [Bronzini, 2007]. Here, there is a specialised workforce. The [Bronzini, 2007]'s paper evidences statistical significance of location externalities on FDIs.

These concepts have been introduced to understand Italian production performance. Through direct investment flows, an investor builds up a FDI position that has an impact on an economy's international investment position. This FDI position (or FDI stock) differs from the accumulated flows because of revaluation (changes in prices or exchange rates) and other adjustments like rescheduling or cancellation of loans or debt-equity swaps [Eurostat, 2018]. As noted in [Bank of Italy, 2020] report, Italy's annual inward FDI stocks are classified by immediate counterpart country (IMC) and instrument. The total amount of annual inward FDI stocks are 382.427 million euros. The applied principle for estimating the FDI is the extended directional principle differs from the asset/liability principle used for aggregate direct investment statistics that are included in balance of payments and international investment position accounts. The difference between the two principles depends on the treatment of reverse investment (i.e. when an affiliate invests in its parent). Under the extended directional presentation, reverse investment is subtracted to derive the amount of total outward investment of the reporting country: if a resident parent borrows money from one of its foreign affiliates, this is subtracted in calculating the reporting country's outward investment because it reduces the amount of money that country's parents have invested in their foreign affiliates. Under the asset/liability presentation, reverse investment is instead recorded as a liability of the reporting country [Bank of Italy, 2020].

2.3 Total Factory Productivity

Total Factor Productivity (TFP) is the portion of output not explained by the number of inputs used in production. As such [Comin, 2010], its level is determined by how efficiently and intensely the inputs are utilized in the production. Usually, the growth rate of TFP is measured by the Solow residual, measured by $g_y - \alpha^* g_k - (1-\alpha)^* g_l$; where g_y is the growth rate of aggregate output, g_k is the growth rate of aggregate capital, g_l is the growth rate of aggregate labour under the assumptions of a) the production function is Cobb-Douglass, b) there is perfect competition in the factor market, and c) the growth rates of output and the inputs are measured accurately [Comin, 2010].

TFP is a measure of growth proper than Gross Domestic Product (GDP), and it is technologically-driven advancement through technical change [Hulten, 2001]. It is widely used for R&D models even if quantifying R&D spending is difficult because essentially an internal investment to the firm, with no observable "asset" price associated with the investment "good" and no observable income, streams with the stock of R&D capital [Hulten, 2001].

The interest in R&D as an endogenous explanation of output growth and innovation is recognized as a form of capital accumulation. R&D and human capital are determinants of outputs growth [Hulten, 2001]. It is important to note that GDP and TFP are complementary, not substitutes for measuring the growth rate and the effect of innovation on society's welfare [Hulten, 2001]. It is important to introduce these concepts for our analysis. For example, in Italy, a gap is evident between regions to the specific role of absorptive capacity, technological capabilities, diffusion and knowledge spillovers. It also shows how important is the role of regional policies and spatial strategies in balancing and equalizing the disparities due to private capital and public infrastructure investments from rapidly expanding developed areas to under-developed areas [Napolitano et al., 2018].

As FRED [University of Groningen and University of California, Davis, 2022a] [University of Groningen and University of California, Davis, 2022b], the real GDP is increasing and TFP is decreasing, as it possible to see in the graph:

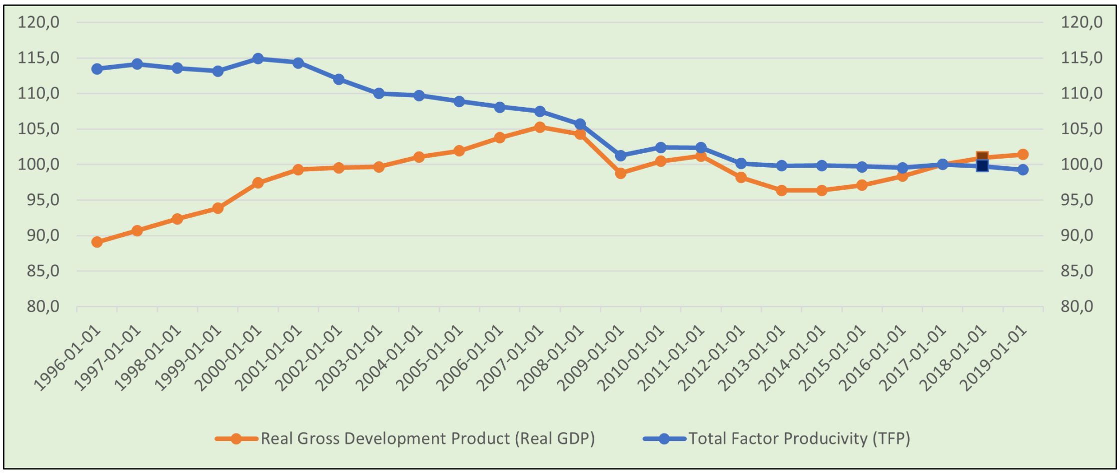

FIGURE 2.1: Annual trend not seasonally adjusted of Real GDP from 1995 to 2019, index 2017=100 and annual trend not seasonally adjusted of TFP from 1995 to 2019 (at constant national prices), index 2017=100 - Our elaboration

It describes the annual trend not seasonally adjusted of real GDP and TFP at constant national prices from 1995 to 2019, 2017 has been assumed to be the index and is equal to 100, indeed it is marked differently.

As [Garofalo and Pugliesi, 2014], TFP is the main driver of Italian growth during the 1890s and 1990s. Starting from mid-1990s, it starts to slow down due to macro factors such as low ICT and low internationalization, micro factors such as industry specification and firm dimension, institutional factor such as ownership structures of enterprises and approach and innovation policies, and other factors related to the job market and its policies.

2.4 Strengths and Weaknesses of the Italian Production System

The Italian production system is mainly small-medium enterprise based. There are some restrictions to the development of the Italian industry, for example, the high bureaucratization rate and high taxation rate. In addition, Italian industry is characterized by low value-added activities. There is, indeed, a growth delay because there are continuous changes in the regulatory environment.

The Italian market is more oriented toward a monopolistic market as there is not sufficient enforcement to promote competition and a modern market economy. Overregulation, which is anti-competitive, produces a negative effect on economic performance. There are strong anti-competitive restrictions and high barriers to entry that contribute to protecting the incumbent companies already operating in the market. Exogenous shocks have influenced the Italian economy as the technology revolution, the birth of the IoT sector, globalization, integration of global financial markets and adoption of the euro, and in conclusion, several crises in 2008-2009, 2014-2015 and the Covid-19 crisis have affected the country.

There is a lack of economic growth due to unsolved structural problems although there are several differences in the diverse sectors and the high shadow economy incidence, their massive migration flows, and the intense delocalization process, as cited in [Brandolini and Bugame
report: "it appears difficult to reconcile the strong and steady growth in employment with stagnating production"; there is also lack of foreign capitals due to high taxation.

In Italy, GDP growth is slow because there are a low amount of investments in R&D and labour productivity. Labour productivity is decreasing and therefore raises decreasingly as the quantity produced increases. Italy has many difficulties investing in R&D because small companies have limited negotiating power; they are struggling to achieve liquidity from financial intermediaries. It has no economies of scale and investment costs are very high, the production is low-tech content and easily imitated, which causes strong foreign competition, especially from the competitiveness of east Europe.

It is also significant to consider that small-medium enterprise-based companies are characterised by familiar governance; it obstructs private equity and the entrance into the stock market. As far as labour productivity, it is possible to note that there is a human capital specialization lack due to issues in Italian education. Labour productivity is characterized by instability and too much term workforce [Brandolini and Bugamelli, 2009]. For example, in the agri-food and made-in-Italy sectors, investments in R&D are lower because they are activities with low value-added [Brandolini and Bugamelli, 2009].
As noted in [Cassa Depositi e Prestiti, 2018], the last twenty years productivity is stagnant, and the divergence with the rest of Europe is increased. Because of the above, the R&D investments are only 1.51% of GDP. Especially in recent years, an effort has been made important in supporting technological development and the modernization of process-production processes [Cassa Depositi e Prestiti, 2018]. The plan for 4.0 is critically significant for the technologies' development in the Italian production system, it is evident moreover that smart specialization diffusion is major in companies operating in design, creativity and Made in Italy, and automotive. It is further evidence that businesses in networks are investing more in R&D.
The Italian production system has not only weaknesses but also strengths. They are greater flexibility in decision-making due to small dimensions. Some strong points in the luxury industry are the high specialization industry, the option to customise purchased products, and a direct relationship between customers and company.
The Italian production is very connected to the territory, as evidenced by [Marshall, 2009], the presence of a cluster of companies specialized in one region creates positive externalities and supports the district foundation [Di Giacinto et al., 2022].
The Italian districts experimented with an important structural change the increasing global participation in the value chain, and the bigger companies belonging to industrial districts have taken on an increasingly important role.

2.5 Differences between Industrial Districts and Business Networks

At this point in the treatment, it is on-topic to define the difference between industrial districts and business networks. [Marshall, 2009] has been the first who has studied industrial districts. They are made up of a group of small companies operating in the same

sectors and local areas. They are very gathered in the territory and dependent on each other. Small businesses cannot take advantage of the internal economy of scale of small firms, but they can exploit external economies of scale that are internal to the industrial districts. These economies are localised and typical of the territory [Garofalo, 2007].
District externalities are path dependence and they are from economic variables connected to the region as the social capital with different aspects: learning by interacting, accessing public assets, local rivalry, and peer pressure.
Law no. 317/1991 recognises industrial districts as measures for the innovation and development of small and medium-sized enterprises [Garofalo, 2007].
In districts, firms collaborate and create their relationship networks and connections. This phenomenon is grown across the country, and in 2009 they are formalized in business networks [Burlina, 2020]. Business networks are not constrained within territorial boundaries but they may exist networks inter-region and inter-sector [Burlina, 2020]. It is likely to consider business networks as an expansion of industrial districts, indeed it is no accident that most business networks are even concentrated in the same province. It is possible to note that informal collaborations are long-standing, and then in 2009 they were formalised. Both organizational forms want to fight the constraints connected with the small dimension of firms. Several limits forbid the growth made by R&D furthermore industrial districts have some difficulties accessing funding [Garofalo, 2007], and business networks want to facilitate it.
There are information asymmetries in both of them and efficient economic policies aiming to reduce asymmetries do not exist.
Since globalization has negatively affected the Italian production system, small-medium firms cannot compete in an international market. Districts and networks aim to increment productivity and international competitiveness. As mentioned in [Burlina, 2020], networks allow firms are not geographically proximate to cooperate and develop specific economic projects; they contribute to producing knowledge spillover and new ideas sharing. There is some evidence that networks improve firms' economic performance.
As noted in [Zecca et al., 2014], industrial districts are reference's territorial unit and network contracts, that are less expensive and easier to be established, must place side by side to integrate and eventually substitute functionalities.

2.6 Importance of Business Networks

Business networks are strictly important for the Italian industry growth. They are a different suitable unicum for the Italian industry [Cabigiosu, 2021a]. The phenomenon is increased and it has assumed a macro-economic relevance.
They enable SMEs firms to develop R&D, take advantage of economies of scope and economies of scale, improve efficiency and relationship with financial intermediaries, increase revenues, improve their competitiveness, and operate in markets with medium-large enterprise [Pastore et al., 2020].
This organizational form fits well with the Italian system, based on SMEs.
In the literature there is moreover evidence that cooperation and knowledge transfer affect positively R&D: quality effort (product innovation) decreases with the number of connections whereas increases in presence of process innovation. Quality-improving networks are denser relative to those driven by process innovation, link formation is welfare improving if both absorptive capacity and quality spillovers are sufficiently small [Di Dio and Correani, 2020]. These aspects contribute to affect welfare [Di Dio and Correani, 2020].
R&D needs financial capitals sometimes not available for SMEs, and banks require more guarantees, it is clearly that SMEs operating in networks achieve capital because there is a condition's improvement of access to bank credit, both qualitatively (more favorable rating) and quantitative (greater credit line granted and economic conditions facilitated) as networks have a bigger negotiating power. In addition, networks may exploit funds reserved for business networks and SMEs [Proto, 2021].
It is possible to note that in "reti-soggetto" raising revenues are bigger as shown by ROS index, providing evidence of firms' operational efficiency; ROE index, a factor that could improve the ability of firms to attract investors; liquidity ratio, that has consequent beneficial effects on firms' solvency and capacity to repay the debt; and debt's sustainability; this result is achieved in part due to the growth of Ebitda [Pastore et al., 2020].
This growth contributes to producing market externalities, independently from the market structure [Economides, 1996], and there is more propensity for networking, technological innovation and sharing of ideas and knowledge in companies belonging to a business network contract.
There is indeed, a correlation between network membership's growth and the increasing of the profitability of its member' enterprises as confirmed by [Cabigiosu, 2021b]. Because of this, a virtuous cycle is established and it is possible to say that business

networks can be one of the solutions to the Italian problem connected to the industry. The instrument needs more incentives and development by the economic actors as policy maker and financial intermediaries.

2.7 Characteristics and Impact of Firms in Networks

Firms operating in business networks have some similarities. First of all, it is relevant to say that most enterprises members of the same business network are geographically very close. Indeed, 58.4% of them are in the same province, and belongs to different sectors [Romano et al., 2016]. Networks are composed of a small group of business companies. Two important drivers of these companies are the openness to the foreign market and international investment in R&D. They are the goal of the networks' contract foundation.

Looking at the legal form, most companies are Public Limited Companies (PLCs). Many of them operate in the agricultural sector. The 45.8% is composed of micro-companies, which stand for less than 10 employees and small companies, from 10 to 49 employees; but the newest entrant companies are much bigger than the companies dimension in the Italian industry [Romano et al., 2016]. Increasing the size of companies, increasing the business capacity to innovate and use economies of scale and scope.

As [Romano et al., 2016], the connectivity rate with the rest of the productive system is higher with a gap at each dimensional level, except for the biggest ones. There are several possibilities for companies to establish links with other organizations, in truth network contracts are complementary with other business forms of companies' collaborations such as joint ventures, consortia, informal arrangements, and all functional forms for enterprises of small size.

Although it is true that networks companies' members are more efficient with a high added value. Desegregation by macro-sectors of activity also indicates that the higher average productivity of networked firms mainly concerns the manufacturing, while in market services the difference compared to non-in-the-network is not significant [Romano et al., 2016]. They are more inclined to geographically distant markets and to innovation, business networks are born to encourage research and technological innovation [Romano et al., 2016]. In the last years, companies joining in-network contracts are bigger [Romano et al., 2016]. It is important to bear in mind that companies that participate in business networks were

already more inclined to the foreign market and competitiveness [Costa et al., 2017]. Network agreement has had cumulative effects and entering into a network facilitates entrance into other network contracts [Costa et al., 2017]. It is possible to see that network contracts have preserved the competitiveness's enterprises during the crisis. As in [Costa et al., 2017] the companies' revenues increased significantly and the development of employees and revenue has had a different impact according to the production sector [Costa et al., 2017].
In particular, the average turnover has doubled for micro-enterprises after the first three years [Costa et al., 2017]. There are only positive effects for companies into networks, furthermore the impact on trust financial intermediaries to fulfil their obligations [Costa et al., 2017].
In conclusion, network contracts are convenient for companies, they allow to overpass geographical distance, sector and dimension constraints of each company, and they exhibit as a valid industrial policy instrument available to policymakers.

2.8 Description of Networks

It is possible to distinguish two forms of network contracts: "reti-soggetto" and, "reti-contratto". Both are constituted by public deeds and have their tax code. Formal business collaborations should also have a VAT number, name and location of the network, an agency and a common fund, instead of informal collaborations. Formal business networks must pay the annual chamber fee and file the balance sheet. "Reti-contratto" can get tax breaks and have fewer tax and legal requirements than formal collaborations, in a matter of fact they have not got legal subjectivity and they are more similar to a pure contract. In July 2022, 85% of networks are informal collaborations. In total, the number of business networks is 7839 [Camere di Commercio d'Italia, 2022].

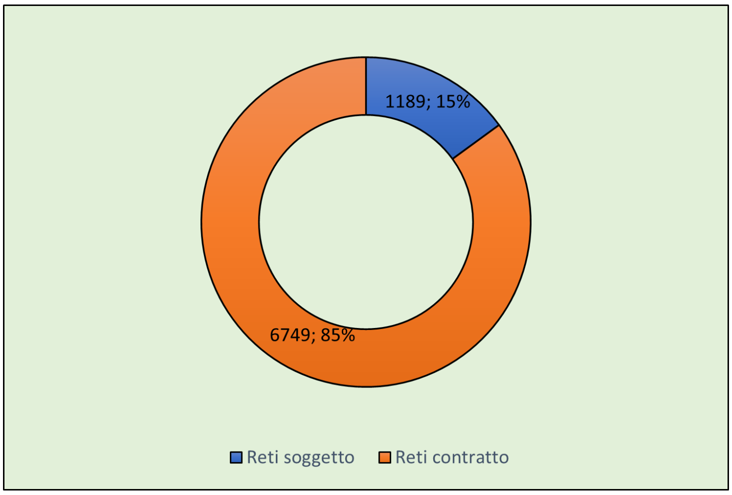

FIGURE 2.2: Type of Business Network Agreement in 2022 - Our elaboration

This phenomenon increased over the years, and it raised a macro-economic relevance. In the graphic, the increasing trend is shown.

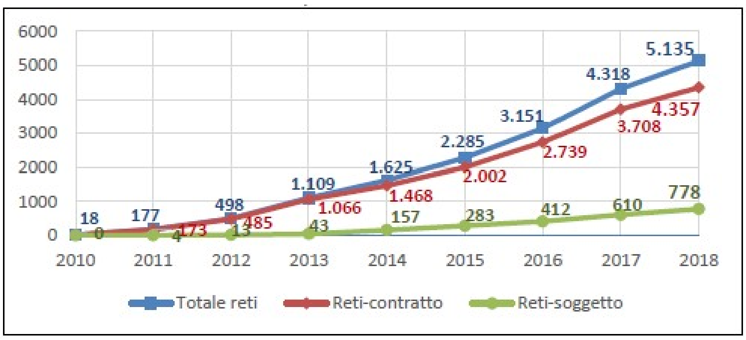

FIGURE 2.3: Trend of network contracts, 2010-2018 [RetImpresa, 2019b]

2.8.1 Date of Network Foundation

The network contracts were born in 2010. The birth of network contracts is related to laws, and some important tax breaks given by the regions [Bartoli and Rizzi, 2017]. In fact, the regions give financial contributions to the networks from 2010 to 2016. In total, the number of funds granted in this period is equal to 1,1 billion euro, compared to 2,2 billion euro allocated. In light blue, it is possible to see the number of funds provided to the network' projects.

The goals of these funds are R&D, business development, internationalization, network creation, environmental investment and welfare. Since 2013 tax-level incentives on networks have disappeared (L.122/2010).

RetImpresa aligns business networks as a valid instrument of business economic development and a collaboration's model for the country. The trend is also influenced by regions and other institutions, that should collaborate and simplify the relationship with the enterprise and networks.

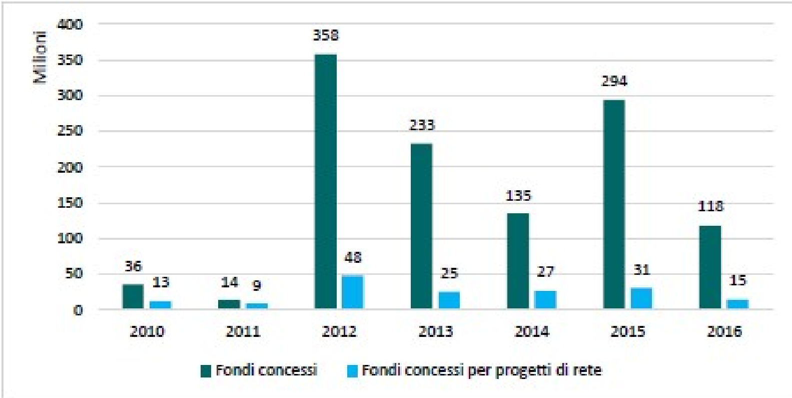

FIGURE 2.4: Funds granted to firms and networks during 2010-2016, [Bartoli and Rizzi, 2017]

Here, there is a monthly and annual trend of the foundation of "Reti-contratto". Both graphs report the time on the x-axis, one given in months, the other ones in years, and on the y-axis the number of contracts.

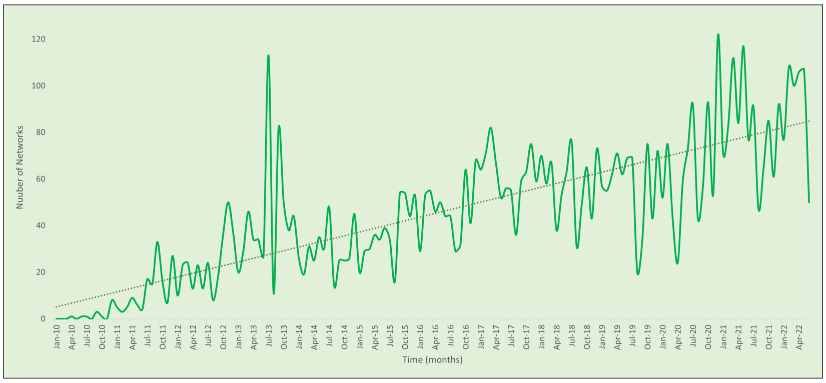

FIGURE 2.5: Monthly Trend of Network Foundation - Our elaboration

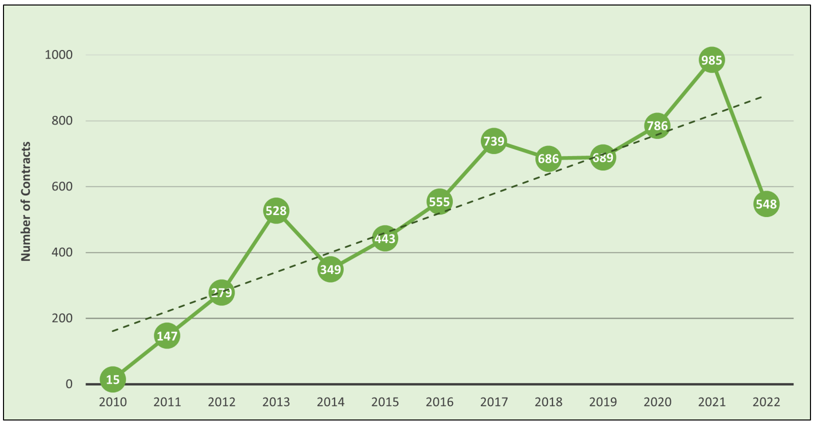

FIGURE 2.6: Annual Trend of Networks Foundation - Our elaboration

In both cases, the trend is increasing over the years. In the first chart, there is a peak in the quarter from April to July 2013 and another one in the quarter from October 2020 to January 2021.

In the annual trend, there is an interesting frequency of cyclic events of the establishment of network contracts because every four years, there is a slight increment of the foundation of network contracts. Unfortunately, as the youth of the phenomenon, there are no other data to understand whether this behaviour is cyclical or not.

2.8.2 Network's Structure

Social Networks Analysis (SNA) is employed in the economy to investigate social structure through the use of networks and graph theory. It allows analysing the single firms belonging to a network and the effect related to the structure of networks [Sciarelli and Tani, 2014]. Formally, networks are composed of links that connect nodes, representing the single firms [Economides, 1996].

In our analysis, we consider a fully connected network because each company is connected with N-1 firms in the network. All members have direct ties to each other. The network's density is a good proxy to evaluate the presence of deterrence-based trust. The study of the network's structure is very helpful to measure the network's innovative capacity [Sciarelli and Tani, 2014].

It is meaningful to consider that a dense web of interaction in the core of the network with selective relations to the network periphery is assumed to guarantee efficient knowledge and information diffusion throughout the network's structure [Wanzenböck, 2018]. As in our analysis, it is possible to sustain that efficient knowledge and information diffusion affect positively the productivity of the network.

Network proximity is a good indicator of how an actor of each node, that corresponds to each firm, is close or far away from the elements of the network's structure. Measuring the network's proximity of aggregates may be a valuable extension of the common relationship, distance or closeness measures in SNA [Wanzenböck, 2018]. The mentioned article measures the proximity across regions. In this study, the proximity has been measured across firms in the same network to try to understand if linkages between firms affect the type of goods produced, the dynamism of the network, and the production of positive externalities.

The produced externalities may be lightly different, given the two types of business networks: horizontal and vertical. The horizontal business network is a network of companies operating in the same market unlike the vertical one is a network of companies operating in different stages of the production chain.

Density measures the network's connection, standing for the average number of partners for each firm, and dynamism measure the ability to attract and accept new partners. These characteristics influence competitive and innovative performance. A very dense network with strong and cohesive ties allows knowledge transfer, and the opening to new market gives the possibility to enter a foreign markets and new technologies, producing new goods. Networks can be descripted as a "small world" in which the company exploits synergies, economies of scale, variety and internal cohesion [Garofalo and Pugliesi, 2014]. These network characteristics are analyzed in the following sections. They are the degree of connection, which correspond to the number of nodes in each network, the territorial extension to establish if the network is local or cross-territorial, and the economic range, which is the number of affected macro-area of activity [Garofalo and Pugliesi, 2014].

2.8.3 Average Number

As mentioned in [Pitingaro and Corsini, 2021], whole network aggregations registered in mid-2021 are divided as follows, 86% are composed under ten enterprises, and more than 50% consist of micro-aggregations (2-3 enterprises). The data show polarization of network contracts on the extremes of entrepreneurial density, with a strengthening of micro-networks and a slight decrease in all other sizes; in most cases, the companies belong to the same region. It is also important to observe the following graph:

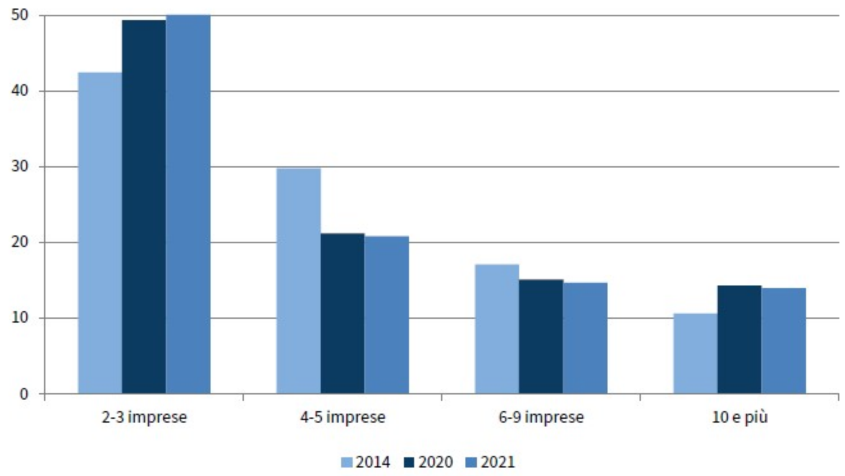

FIGURE 2.7: Network contracts by no. of firms involved (% of total), [Pitingaro and Corsini, 2021]

It can be seen that in 2021, networks composed of 2-3 firms increased, and those composed of more than ten networks decreased concerning the previous year. It shows that the network's dimension is raised for the smallest and the biggest in 2014. More updated results are in the following chapter.

2.8.4 Regions

The firms belonging to network contracts are placed throughout the Italian territory, as shown in the table.

Regions	**Numbers of Companies**
Abruzzo	1395
Basilicata	416
Calabria	882
Campania	3298
Emilia-Romagna	2601
Friuli-Venezia Giulia	2264
Lazio	9938
Liguria	1066
Lombardia	4565
Marche	1263
Molise	118
Piemonte	2183
Puglia	2493
Sardegna	1054
Sicilia	1527
Toscana	3142
Trenino-Alto Adige	756
Umbria	1033
Valle d'Aosta	159
Veneto	3562
TOTAL	43715

TABLE 2.1: Companies involved in business networks by region [Camere di Commercio d'Italia, 2022]

Lazio, Lombardia, Campania, Veneto, and Puglia are the regions with the greatest number of companies involved in business networks.

As in [RetImpresa, 2019b], the propensity to networking is higher in Lazio, Friuli-Venezia and Umbria. It shows that in these areas, it is present more incentive for networking. Propensity to networking is higher in Friuli Venezia-Giulia and Lazio, but the propensity to networking seems to be strengthening year by year and the trend to networking is increasing [RetImpresa, 2021]. In the following pie, it is possible to see the division of network contracts by the North, Middle, and South of Italy. This chart suggests that in the south the productive environment is less developed.

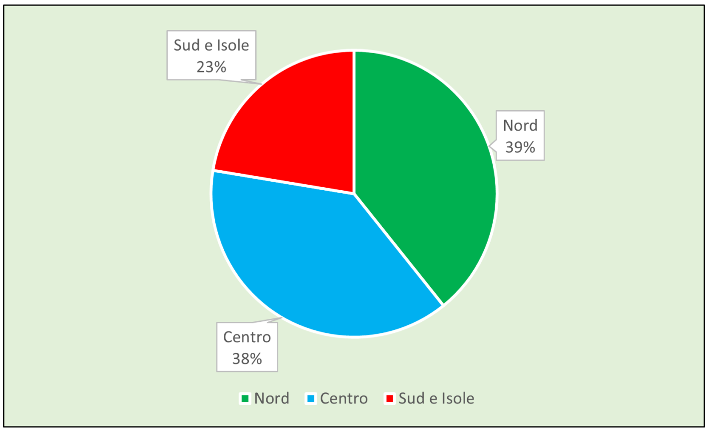

FIGURE 2.8: Pie chart of business networks by geographic area - Our elaboration

As [RetImpresa, 2021], the 75% of networks is single-regional, and 25% is multi-regional; this consideration is important to understand the birth of network contracts in Italy, [Zazzaro, 2010] sustains that most of the network contracts were born in the same industrial district and they are the precursors of the networks.

2.8.5 Macro-Area of Activities

Macro-Area of activities is a grouping of different specific sectors that operate in the same field, known as the production sector. The production sector is the industrial sector in which a firm operates. The classification of the Ateco code by [Istat, 2022b] was used by [Pitingaro and Corsini, 2021] to classify companies. In 2019, agriculture, commerce, manufacturing, and tourism services were the most frequent [RetImpresa, 2019a]. In 2020, agriculture, construction, commerce, and mechanical industry were the production sectors with the higher value of percentage distribution of networks [RetImpresa, 2020], and in 2021 we have 22,2% firms in agriculture, 14,4% in commerce, 12.1% in construction, and 10% in tourism [Pitingaro and Corsini, 2021].

Now agriculture, constructions, commerce, and mechanics are the sectors with the higher number of companies participating in a network contract. In July 2022, it is evident that there is an increasing trend of network contracts of firms that operate in the agriculture, fashion and furniture system, unfortunately, there are more than 20% of sector values that are missing.

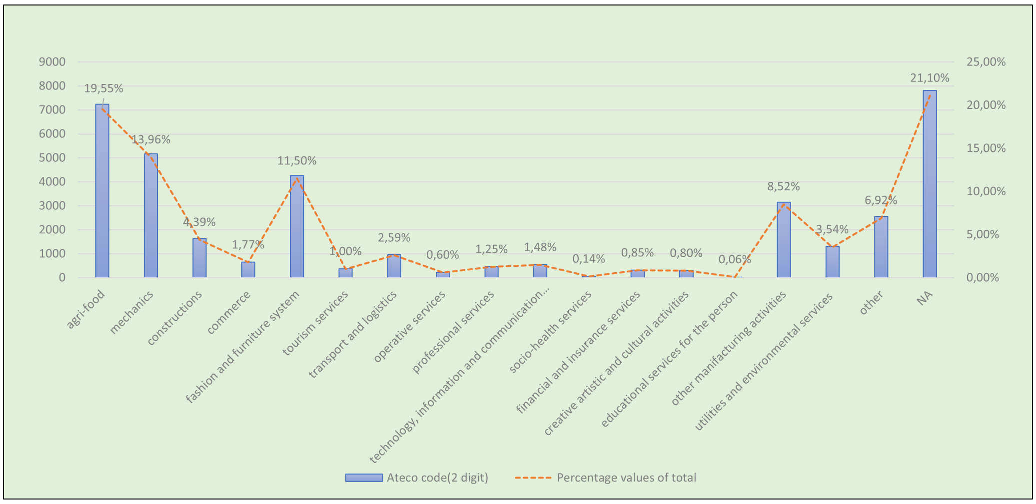

FIGURE 2.9: Macro-area of activity of companies belonging to networks - Our elaboration

Agriculture, fashion and furniture system, and mechanics are the kinds of macro-area sectors that are higher. For this figure, it was used the same classification as the others report [RetImpresa, 2019b, RetImpresa, 2020, Pitingaro and Corsini, 2021]. In the appendix, it is possible to see the classification criteria.

2.8.6 Business Sector

Following the classification of Ateco code by [Istat, 2022b], the following results were obtained:

A 02	184	C 21	25	E 36	35	H 53	116	M 70	954	Q 86	1144
A 03	131	C 22	210	E 37	19	I 55	1085	M 71	515	Q 87	161
B 08	33	C 23	219	E 38	226	I 56	1195	M 72	258	Q 88	354
B 09	1	C 24	69	E 39	15	J 58	88	M 73	247	R 90	95
C 10	916	C 25	1145	F 41	1276	J 59	63	M 74	391	R 91	58
C 11	194	C 26	208	F 42	254	J 60	22	M 75	6	R 92	16
C 13	257	C 27	201	F 43	2571	J 61	55	N 77	157	R 93	316
C 14	170	C 28	554	G 45	341	J 62	1025	N 78	49	r NA	1
C 15	275	C 29	91	G 46	1343	J 63	326	N 79	439	S 94	5
C 16	229	C 30	85	G 47	1658	K 64	255	N 80	111	S 95	69
C 17	50	C 31	264	H 49	1136	K 65	5	N 81	697	S 96	362
C 18	107	C 32	158	H 50	51	K 66	149	N 82	857		

FIGURE 2.10: Distribution of Ateco code (2 Digit) - Our elaboration

Highest values are obtained in A1, A2, A3 (corresponding to agriculture sector), B8 B9, C25, C10, and C28.

Ateco code is the classification of the economic activities by [Istat, 2022b] for statistical purposes. This kind of classification is used at the national and international levels. In Italy, the set is employed for administrative and fiscal topics. The Ateco code doesn't have legal value, but it is used when a VAT number is registered at the administration. The first letter indicates the section i.e. the activity in general, and the numbers indicates the division, i.e. more detailed activity [Istat, 2022a].

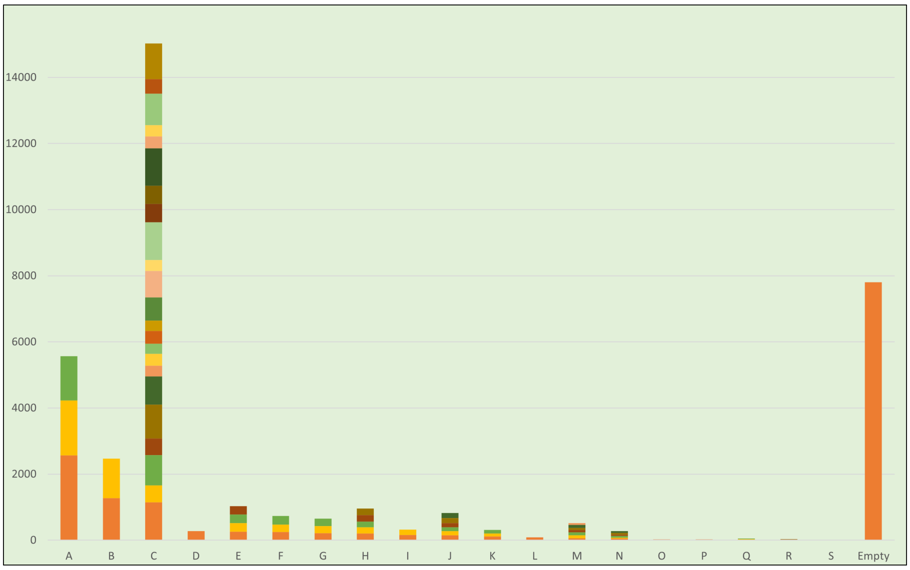

FIGURE 2.11: Activity section of the enterprises belonging to the network - Our elaboration

This chart shows another classification criterion, bars indicate the section, and the different colours indicate divisions. Other more detailed classifications are overlooked. Figure 2.11 shows that most firms work in manufacturing activities (C) and agriculture sectors (A). All data are consistent with each other.

In the next chapter empirical model is displayed. Regions, business sectors, and average numbers are explained by the network at an aggregate level.

Chapter 3

Empirical Data and Related Literature

The goal of this chapter is to apply a regression model and investigate if the results support the evidence provided by [Di Dio et al., 2022] model.
The chapter is organised into five sections. In the first one, the theoretical background is illustrated. The second section explains the dataset and its features. The third paragraph details the variables of the econometric regression which are the number of companies, regions, sectors, and types of goods computed with entropy, and the average distance. The paragraphs conclude with the empirical model and results.
All the data have been analysed with Excel and R 4.2.2 software, and then collected in a unique CSV file.

3.1 Theoretical Model

For our analysis, it is necessary to introduce the [Di Dio et al., 2022] model which is mainly the work on which our analysis is based.
The paper contributes to the stand of literature addressing network formation and R&D agreements among vertically and horizontally differentiated firms. It tries answer to the following question: how networks among firms can channel R&D investments towards improvements in productivity and social welfare?
The model tries to investigate the relationship between the nature of goods produced

by firms and their spatial dispersion within an already-formed and stable network. Furthermore, an oligopolistic competition Cournòt model is developed and can account if business networks are formed independently of the nature of goods produced by participating firms and if the nature of produced goods affects the distance between firms of the same networks.
The model is solved as a tree-stage game with imperfect information: in the first stage, firms decide whether to form a collaborative link; in the second one, they choose their location; and in the final stage, firms invest in cost-reducing technology simultaneously indeed, investment and production are considered.
Every firm chooses R&D effort unilaterally; it allows it to reduce its own marginal cost of production, but it also has positive spillover on the cost of the rival firms, endogenous spillovers (or absorptive capacity) are assumed.
The model is the result of the combination of an oligopolistic competition a là Cournòt with the typical structure of the Salop-Pal model where every point represents both a location and a market. It considers the real situation of firms producing homogeneous goods but located at different points of the Salop circle, so it is possible to affirm that the greater the distance between the companies, the higher the product differentiation. It is important to reiterate that when one firm invests, it is not able to observe both the production and the investment of a rival firm.
In summary, strategic location, R&D investment, and differentiated production are taken into consideration.

The focal point of the model is shown analytically by the following equation:

$$\pi_i = [A - q_i + Eq_j - t\left|x - x_i\right| - c_i]q_i - \gamma\frac{z_i^2}{2} \forall i \neq j \in 1,2$$

where π_i is the profit of firm i, A indicates the market size, q_i is the quantity of goods produced by firm i, E considers the product differentiation varying from [-1,0) if the goods are substitute, E=0 if the goods are independent and E=(0,1) if goods are complement (in the case E=1 goods are perfect complements), $t\left|x - x_i\right|$ is the transport cost with $t > 0$; c_i is the marginal cost of firm i; z_i indicates firm i's R&D effort and γ is his parameter.
In the first stage, firms decide whether to form a collaborative link or not. The link is formed only if both firms agree to it. In the second stage, they choose their location x_1 and x_2 on Salop's circle with circumference 1. In this case, to consider imperfect

information of firms, investment and production are simultaneous.
The model's key findings are the following:

- first, two firms will tend to form stable cooperative links irrespective of the nature of the goods they produce,
- and second, the distance between the two cooperative firms will tend to be high in the case of substitute goods and lower agglomeration in the case of complementary or independent goods.

R&D effort is affected by location even when firms are monopolists. They will tend to agglomerate since the smaller the distance from the rival firm the higher the positive effect (externalities) from R&D cooperation indeed, R&D investment in a specific market x tends to decrease as the distance from this market increases. It is also shown that the cooperative link is stable regardless of the type of goods produced and the location choice. It is important to point out that when goods are perfect substitutes, a closed position of firm j reduces the firm i's quantity and so profits, pushing firm i to locate the maximum distance from j. Similar opposite considerations are valid when goods are the perfect complement.
Moreover, cooperation allows monopolistic firms to take advantage of positive spillover absorbing the other firm's R&D investment without any risk of encouraging opportunistic behaviour by rival companies, reducing marginal costs and increasing production and finally profits.
The paper's empirical evidence shows that the average distance between companies in the same network tends to be greater in the case in which participating firms produced substitute goods/services, thus providing further evidence about the role of the nature of goods produced on the distance between firms in forming links.
In a nutshell, the nature of goods produced decisively influences the degree of network dispersion, as cooperative firms cooperate will agglomerate in the case of complementary goods whereas they will disperse if goods are substitutes. Differently, cooperation implies agglomeration if firms produce independent goods and spillover is strictly positive.
The goal of this thesis aims to investigate with empirical data [Camere di Commercio d'Italia, 2022] if the [Di Dio et al., 2022] model is supported by empirical evidence. In the next sections, an empirical model has been developed to achieve this purpose. The dataset, variables, model and results are illustrated.

3.2 Dataset

This paragraph aims to describe the dataset, taken from [Camere di Commercio d'Italia, 2022] site [1], on which our analysis is based.
The dataset focuses attention on business networks, and here it is possible to note that there are three different sheets. The first one is a summary of the companies belonging to the network's agreements, filtered by regions. In the second sheet, there is the empirical data about "reti-contratto" which are an informal collaboration, and the third one describes formal collaborations, which are "reti-soggetto".
Our dataset is organized in rows and columns. In the first row, there is the header, and in the other 36996 rows, there are details about each company.
By column, the following attributes can be distinguished:

- **Progr.**: Number of network programs which the companies belongs to
- **Denominazione contratto**: Name of network contract
- **Data atto**: Date of the act of foundation
- **Numero repertorio**
- **Numero atto**
- **Oggetto**: Reason why enterprises cooperate
- **N.area**
- **C.F.**: Tax code
- **Denominazione impresa**: Name of the company
- **Impresa di riferimento**
- **Comune**: City where the company is located
- **REG**: Region where the company is located
- **PV**: Province where the company is located
- **NG**

[1] https://contrattidirete.registroimprese.it/reti/

- **Codice ateco 2007**: Ateco code of each company, that is the production sector of the enterprise, numeric version
- **Settore attività**: Industrial sector of the enterprise
- **Sezione attività**: Macro production code of the enterprise
- **Attività**: Description of companies activities using the Ateco code of each company, 2 digit numeric version by [Istat, 2022b]

Unfortunately, there is no metadata describing in more detail the dataset's variables. For our analysis, the following attributes were not considered: *numero repertorio, numero atto, n. area, C.F., impresa di riferimento, NG.*
Most of the variables are qualitative, like *denominazione contratto, oggetto, n. area, C.F., denominazione impresa, impresa di riferimento, comune, PV, NG, settore attività, sezione attività* and *attività*; instead *Progr., data atto, numero repertorio, numero atto, n. area, REG.* are discrete quantitative.

3.3 Variables

In this section, the econometric model's variables are exhibited. They can be classified as independent and dependent variables. The first ones are the number of companies, regions, sectors and types of goods. Instead, the dependent variable is the network's average distance. Intuitively, the number of companies is the variable that describes how many firms are in each network. Regions identify the Italian region where the firms of the networks are mainly based. The sector identifies the production sector in which a firm operates; the last dependent variable is the types of goods that classify the nature of produced goods by networks as independent, complementary or substituted. These variables are related to the average distance, the dependent variable of the empirical model. It computes the average distance between firms belonging to the same network. In the next sub-paragraphs, all the variables are described in more detail.

3.3.1 Number of companies

The number of companies for each network has been computed as in [Pitingaro and Corsini, 2021], and the following results were obtained:

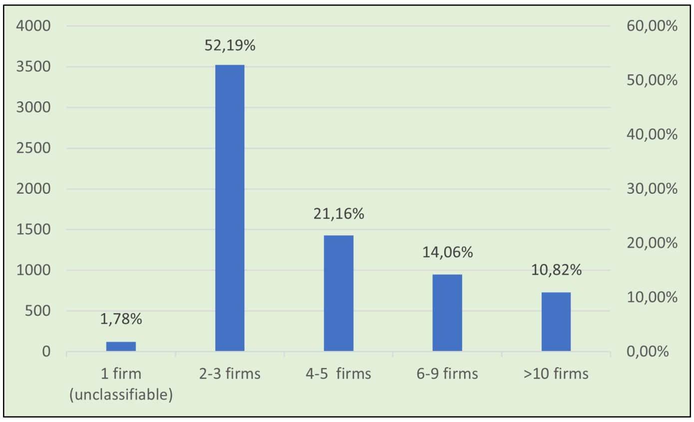

FIGURE 3.1: Number of companies by network - Our elaboration

It is possible to see that the 52.19% of business networks are composed of 2-3 companies; this data seems to be growing to previous years, which might be a little evidence that this kind of collaboration is increasing. In the literature, there is, indeed, evidence that this organizational model of business allows them to overcome the limitation of Small and Medium-sized Enterprises (SMEs) [Pastore et al., 2020]. There are signs that this business model positively affects companies development that belong to the network. It may also influence the regions where the companies and the networks are located: this should be an inspiration for policymakers [Fabrizi et al., 2022]. The 21.16% of the networks is composed of 4-5 firms, as shown the network's dimension is small, mainly micro-based. It is important to note that a firm can take part in more than one network contract, but 91.1% of the total of them belongs to only one network [Pitingaro and Corsini, 2021]. Remarkably, companies belonging to networks are more inclined to agree to other network contracts, accordingly, they are more productive than the others [Romano et al., 2016].

The number of companies in each network is computed as follows: the frequency of the network contract name was calculated. Considering the *Denominazione contratto* and *progr.* attributes, it was counted how many times the same name of network contract appears; the *progr.* identifies the number of the network contract in the data.

3.3.2 Regions

The region variable has been obtained starting from the *REG* variable, a discrete quantitative that indicates from 01 to 20 the Italian regions. Each code stands for an Italian region, as shown in the table:

REG Code	**Region**
01	Piemonte
02	Valle d'Aosta
03	Lombardia
04	Trentino Alto-Adige
05	Veneto
06	Friuli-Venezia Giulia
07	Liguria
08	Emilia-Romagna
09	Toscana
10	Umbria
11	Marche
12	Lazio
13	Abruzzo
14	Molise
15	Campania
16	Puglia
17	Basilicata
18	Calabria
19	Sicilia
20	Sardegna

TABLE 3.1: Region Code [Camere di Commercio d'Italia, 2022]

This variable was reported in each row of the dataset. To obtain the networks' region, the maximum criterion was adopted: it was specified the region code of each network, counting for each network program (provided by *progr.* attribute), the frequency of the code, and assigned to the region that occurs with the maximum frequency; in case of two or more two codes that appear the same number of times, the network was assigned to a multi-region variable that has 21 code. It has repeated the entire process for each network contract to the end of the dataset.

In the end, it has been constructed a dummy variable for each value, taking the value 1 if the network is in the corresponding region, and 0 otherwise.

According to these criteria in the following figures, it is possible to see the territorial distribution of business networks. As shown in both of them, Lombardia and Lazio are the regions with the highest values; in general, regions in the south of Italy have less

number of business network agreements.
It highlights that enterprises are more efficient and productive in the north. The export is superior to the other regions in Lombardia, Veneto, Trentino, Friuli Venezia-Giulia and Lazio [Vendetti, 2022]. This phenomenon is correlated with the southern question, so it contributes to explaining the gap between the north and the south of Italy. Most of the business networks are geographically close [Romano et al., 2016] due to the development of the phenomenon.

Region Name	**Code**	**Number of Business Network Contracts**
Piemonte	01	376
Valle d'Aosta	02	28
Lombardia	03	824
Trentino Alto-Adige	04	123
Veneto	05	684
Friuli Venezia-Giulia	06	340
Liguria	07	123
Emilia-Romagna	08	508
Toscana	09	456
Umbria	10	124
Marche	11	206
Lazio	12	811
Abruzzo	13	275
Molise	14	14
Campania	15	444
Puglia	16	377
Basilicata	17	49
Calabria	18	109
Sicilia	19	210
Sardegna	20	130
Multi region	21	538
Total		**6749**

FIGURE 3.2: Network classification and concentration by region - Our elaboration

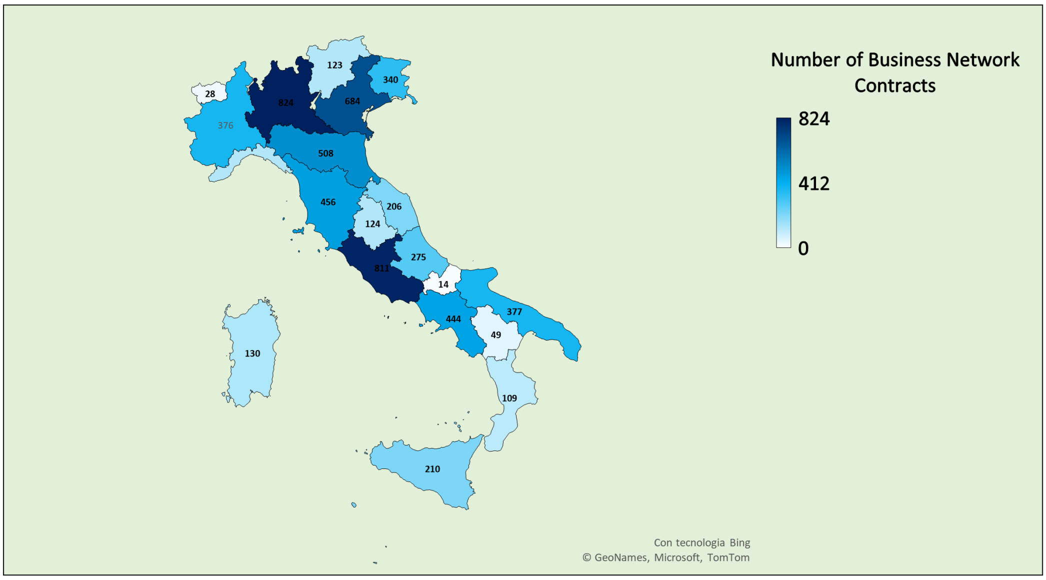

FIGURE 3.3: Geographical representation of business networks - Our elaboration

The regional data were also aggregated according to the NUTS 1 level classification by Eurostat [2]. This classification shows at a more aggregated level the region variable. It is useful to understand all the mentioned phenomena. Indeed, it is evident that in the south and islands, there is only 23% of the total of business network contracts. North-east and middle of Italy are the regions with the higher number of networks. Dummy variables were once again constructed, they consider north-nest, north-east, middle, south, islands and multi-region variables.

They will be included in our analysis through dummy variables with the others considered for the econometric model.

		Percentage Values
North-West	1351	20,02%
North-East	1655	24,52%
Middle	1597	23,66%
South	1268	18,79%
Islands	340	5,04%
Multi-region	538	7,97%
Total	6749	100,00%

FIGURE 3.4: Distribution of business networks by geographical area - Our elaboration

[2] https://ec.europa.eu/eurostat/web/nuts/background

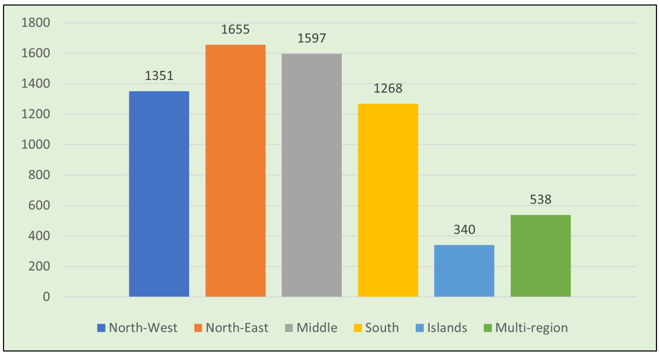

FIGURE 3.5: Histogram of business networks by geographical area- Our elaboration

3.3.3 Sector

Another relevant variable is the business network sector. Starting from the *Settore attività* qualitative variable of our dataset [Camere di Commercio d'Italia, 2022], the network's sector has been computed.
Similarly to the region variable, the maximum criterion was adopted. The different industrial sectors can be grouped in the table:

Sector	**Sector code**
Agricoltura/Pesca	1
Altro Settore	2
Commercio	3
Industria/Artigianato	4
Servizi	5
Turismo	6
Multi-settoriale	7

TABLE 3.2: Sector Code - Our elaboration

Respectively, *Agricoltura/Pesca* stands for Agriculture/Fishing, *Altro Settore* stands for third sectors, *Commercio* is equal to Trade, *Industria/Artigianato* corresponds to Industry/Crafts, *Servizi* is equal to Services, *Turismo* corresponds to Tourism, and *Multi-settoriale* stands for Multi-Sectoral. Here after, English designations will be used. If there are two or more sectors appear with the same frequency, the observation has been classified as multi-sectoral. The third sector indicates the activities typical of non-profits. The entire process was repeated for each network to the end of the dataset. A dummy variable was constructed, 1 if the network belongs to the sector, 0 otherwise. In the following graphs, it is possible to see the distribution of business networks' sector:

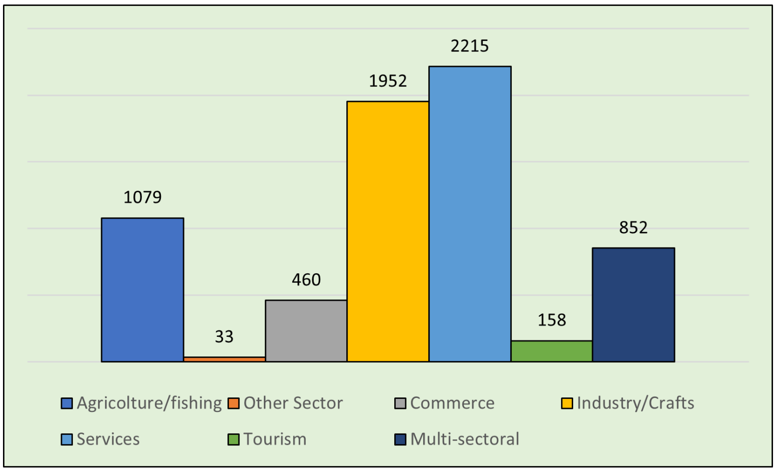

FIGURE 3.6: Histogram of business network's sector - Our elaboration

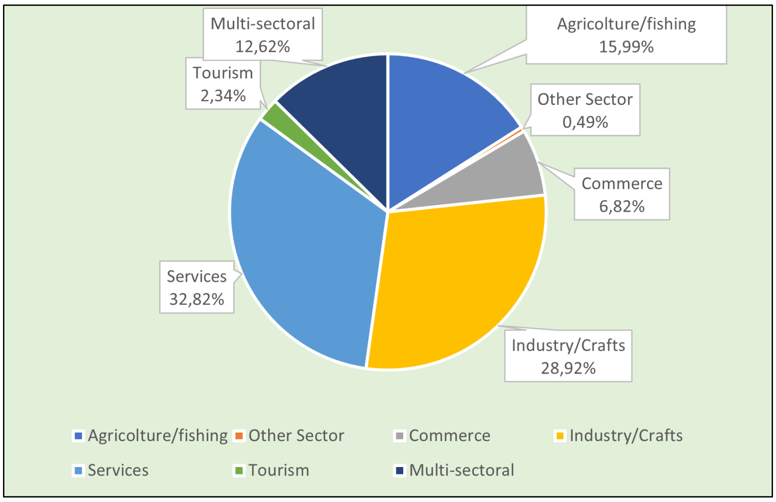

FIGURE 3.7: Pie chart of business network's sector - Our elaboration

Most informal business networks operate in the services (33%) and the industry/craft sector (29%). The services industry is more inclined to establish networks because it is easier to cooperate, despite the distance. Industry/craft is a kind of sector that requires great investments, especially in new equipment, division of the production system's step implies close complementary relations between companies that belong to the same productive sector [Romano et al., 2016].
The amount of percentage of the agriculture sector is considerable. It indicates a low degree of development of Italian firms, with low added value. Participation in partnership by agricultural firms is present in a small number of aggregations [Romano et al., 2016]. It is curious to compare this data with figure 2.9 where the agri-food sector is in greater quantity. Fashion and furniture with other manufacturing services are in decreasing

order: as this classification, there is a reverse situation. The multi-sectoral collaboration is quite frequent. The kind of activity of the firm has to be consistent with the internationalization process, and then firms agree to networks to increment their market power.

3.3.4 Entropy and Type of Goods

Another meaningful independent variable implemented in the model is the type of good or service produced by the network. Entropy has been applied to measure the types of goods or services; in particular, three kinds of goods are possible: complementary, independent or substitute. The variable can take three possible values.
Two goods or services are complementary when the good has to be used with another one to satisfy a consumer's need; for example, a car and gasoline are complementary goods; indeed, they have a related demand and a negative demand cross-elasticity.
Two goods or services are independent when there is no relation between the two goods or services, for example, car and butter; indeed, there is no correlation between their demand functions, and the demand cross-elasticity is equal to 0.
Two goods or services are substituted when the product can be substituted by the consumer with other goods to satisfy equal needs, for example, butter and margarine; in fact, there is a positive demand cross-elasticity.

First of all, it is fundamental to define entropy and try to explain the reason for its implementation.
The *Attività* attribute of the dataset [Camere di Commercio d'Italia, 2022] is composed of text values indicating the sector of activity in which a firm operates following the 2-digit Ateco code version[Istat, 2022b], in addition to a short description of the productive activity was present. Primarily, it was necessary to separate the description and the Ateco code into two distinct parts. Only the Ateco code was used to determine the type of good produced. The variable has been selected after it has been filtered by *progr.* attribute, which indicates the number of business networks in which an organization participates. The entire process was repeated for each network to the end of the dataset. This qualitative mode of character has been used to compute the entropy.
At this point, it is fundamental to define entropy. It is a statistical indicator which measures, in a certain sense, the average information produced by each element in a multiset

[Shannon, 1951], so it is used to determine how the qualitative character mode is distributed among the statistics population [Borra and Di Ciaccio, 2014]. In a frequency distribution, it is possible to have maximum homogeneity when all the units of the statistical collective have the same characterization [Borra and Di Ciaccio, 2014]. Instead, they have maximum heterogeneity when all the units of the statistical collective have different characterization [Borra and Di Ciaccio, 2014]. As [Borra and Di Ciaccio, 2014], in case of maximum homogeneity the sum of the distribution of frequency is equal to 1:

$max \quad homogen.: f_1 = f_2 = ... = f_{j-1} = f_{j+1} = ... = f_K = 0 \quad with \quad f_j = 1$

$max \quad heterogen.: f_1 = f_2 = ... = f_j = ... = f_K = \frac{1}{K}$

Some indicators can measure homogeneity or heterogeneity, one of them is:

$O_2 = \sum_{j=1}^{K} f_j log(f_j)$.

This index is equal to 0 in case of maximum homogeneity. It is also possible to obtain the indicator of entropy from this index:

$Index \quad of \quad Entropy: \quad E_2 = -O_2 = -\sum_{j=1}^{K} f_j log(f_j)$

where $f_i = \frac{n_i}{N}$ is the i-th relative frequency (that corresponds to the proportion in which the character mode is present in the observed collective) [Borra and Di Ciaccio, 2014]. According to the frequentist view and in particular, to [Von Mises, 2014], the probability is the limit to which it tends the event's relative frequency when the number of experiments increases, indeed we have:

$\lim_{n \to \infty} \frac{n_b}{n} = p_b \quad where \quad n_b = n_1 + n_2 + ... + n_j$

This statement is necessary to confirm the equality of the [Shannon, 1951] formula, where entropy is equal to:

$H(X) = -\sum_{i=1}^{N} P(x_i) log_b P(x_i)$

where $P(x_i)$ is the probability that event i occurs, and $log_b P(x_i)$ is the logarithm of the probability that event i occurs and b is the base of the logarithm. The choice of a logarithmic base corresponds to the choice of a unit for measuring information [Shannon, 2001]. To summarize, the entropy of a random variable is the average level of information inherent to the variable's possible outcomes ([Shannon, 1951], [Shannon, 2001]). [Antonietti et al., 2019] applies the entropy information index to measure industry variety which captures the higher probability of a five-digit industry different from an initial distribution being randomly "picked" in a given region; Italian regions vary considerably in skill endowment, performance, and level of development. In this case, entropy has been a measure to classify the type of good produced by networks.

For our investigations, the correspondent values were computed for all the networks of the dataset in more detail the base of the logarithm is assumed to equal 3 because three different kinds of goods are possible.
Entropy has been used as a proxy for the classification of types of produced goods. The criterion is still raw and needs further investigation. For this reason, the variable classification is quite arbitrary, it also depends on the adopted range. As seen, when entropy is equal to 0, the goods are substituted because all the firms produce similar goods, as entropy increases the good becomes firstly complementary, and then independent. The classification leads to three different observed results:

- $H(X) \leq 1$ means that the goods are substituted, and firms operate in the same sectors, so they have the same Ateco codes between them, so the heterogeneity is minimum (or maximum homogeneity), and entropy is equal to 0.
- $1 < H(X) < 2$ means that the goods are complementary, and firms operating in different, but in part, similar sectors, so they have similar Ateco codes between them, heterogeneity is intermediate, and entropy is in the middle.
- $H(X) \geq 2$ means that the goods are independent, and firms operating in different sectors, so they have different Ateco codes between them, so when all the Ateco codes are different the heterogeneity is maximum (or minimum homogeneity), and entropy is greater.

The maximum value, equal to 3.062909, has been obtained by the network program 3307 composed of 63 firms; the minimum value obtained is equal to 0.
As the kind of variable construction, it is possible some subjectivity, sure enough that the adoption of different classification criteria reveals different data.
Another kind of entropy is applied, the normalised version of the parameter, which means that the index changes from 0 to 1. In this case, if the entropy varies from 0 to 0.33 the good is substituted, if it varies from 0.33 to 0.66 it is complementary, and if it varies from 0.66 to 1 it is independent.
The normalization is computed dividing the [Shannon, 1951] entropy by $log_3 n$ [Agarwal and Mittal, 2012].
All the values were computed with the DescTools package [Signorell et mult. al., 2022], for the specific entropy function was used. At aggregate level it is possible to see different

statistics of the type of goods obtained.

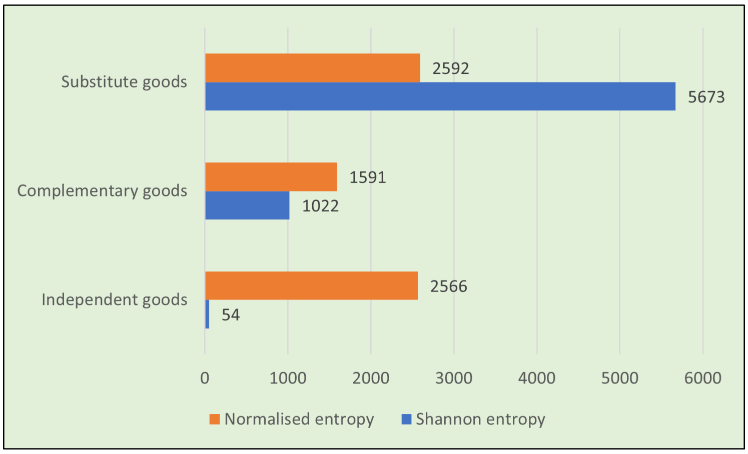

FIGURE 3.8: Different measures of entropy for each type of good - Our elaboration

As shown in the figure, the different kinds of entropy give very significantly different results for each type of good. As per Shannon entropy and normalised criteria, the substituted goods are present in greater quantities, instead for normalised entropy, the independent ones are in larger amounts.

3.3.5 Average Distance

The dependent variable of our model is the average distance. This variable has been computed by the *Comune* attribute indicating the city where the firm is located. Before starting; it is important to point out that another different dataset [3] was used. It contains four attributes: the Istat code, the city's names, the latitude and the longitude. It was merged with [Camere di Commercio d'Italia, 2022] dataset filtering the cities' name that were in common, and another two columns were added to the dataset.

The geographical coordinates were used to compute the network's average euclidean distance. This measurement is a simplification of reality. Euclidean distance was also used for spatial econometric models, as shown in [Buczkowska et al., 2019] where the different methodologies of distance computation are mentioned. For each network, it is possible to identify a matrix (mxn):

[3] https://github.com/MatteoHenryChinaski/Comuni-Italiani-2018-Sql-Json-excel

$$\begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}$$

where x_{11} indicates the distance between the firms with oneself that is equal to 0, but x_{21}, for example, indicates the distance between firms 2 and firm 1 of the same network. The distance between firm 1 and firm 2 is computed intuitively as the distance between two points.

The distance was computed with sf R package [Pebesma, 2018]. The st_as_sf and st_distance commands are applied; first of all, it was necessary to convert coordinates in the geographic distance, and then compute the distance between two points. Firms were selected by *progr.* attribute that indicates the number of network program to which the firms belongs to. Firstly, the distance was expressed in metres, then it has been converted into kilometres. All the results were collected in a matrix.

It is feasible to note that the matrix is always squared. This formula computed the average distance:

$\bar{x} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}}{n(n-1)}$

where n indicates the element of a matrix, this formula corresponds to the correct average distance; the principal diagonal was constituted by all 0.

The total average distance between networks is equal to 70 km, but the quantity of variance is large (14960 km), and the maximum distance is equal to 1003.4 km, instead, the minimum is equal to 0, standing for the firms located in the same city.

Distance is a very relevant variable that may influence the productivity of a network. Indeed, there are some advantages of firms that operate geographically close, as in [Cerutti and De Falco, 2020] because there is the possibility of birth of know-how and culture. Network proximity positively affects business performance, and efficient and fruitful international collaborations are substantively determined by intercultural compatibility, shared values, morale, decision-making structures and objectives, all of which stand for key factors for mutually rewarding partnerships, global mindset and physical distance [Vătămănescu et al., 2020]. As [Vătămănescu et al., 2020], European SMEs are influenced by the gradual internationalization of the organization according to geographic proximity and cross-cultural compatibility.

Distance between firms is in correspondence with major traffic nodes; firms tend to spring

near key junctions of highway networks or large railroad stations [Mori and Nishikimi, 2002], the link between firms and infrastructure defines the density of economies. It is said to be larger if the transport cost per unit of the product is smaller for a given transport density [Mori and Nishikimi, 2002].
Density economies are external to each firm [Mori and Nishikimi, 2002]. Inter-regional transport networks endogenously are determined in the presence of economies of transport density; there is a positive circular causation between them, where there is an inter-regional transport hub as an industrial centre [Mori and Nishikimi, 2002].

3.4 Model

At this point, the econometric model [4] is exhibited:
$distance = \alpha + \beta_1 n + \beta_2 H + \beta_3 sector + \beta_4 area + \epsilon$
where β_i are the correspondent parameters of each regressors, n indicates the number of firms, and H identifies the entropy that has been used to classify the nature of goods: in case H=0 they are perfectly substituted, and to increase the entropy they are independent. The other two are dummies variables: the first one indicates the production sector in which seven dummies are identified, and the other one where the network location according to the NUTS1 level classification by Eurostat [5]. In more detail for the sector: seven dummies are identified: agriculture/fishing, third sector, trade, services, tourism and multi-sectoral, this last variable has been assumed as a control variable.
For the NUTS1 level area: five dummies are distinguished: north-west, north-east, middle, south, islands and multi-region, the latter one has been adopted as a control variable.
The model has been applied to two different datasets, the first one considers only networks with two firms, the small network as possible; the other one considers all the Italian business networks.
A different more detailed version of the model has been applied, because it evaluates not the NUTS1 classification, but the Italian regions. Here it is possible to consider the mentioned model:
$distance = \alpha + \beta_1 n + \beta_2 H + \beta_3 sector + \beta_4 region + \epsilon$

[4] All variables, regressions and related codes can be found at https://github.com/Vcecca/Business-networks. To access contact https://github.com/Vcecca.

[5] The NUTS1 level classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the EU and the UK for socio-economic analyses of the regions at major socio-economic regions https://ec.europa.eu/eurostat/web/nuts/background

where β_i are the regressors' parameter, and n indicates the number of firms, and H identifies the entropy that has been used to classify the nature of goods, in case H=0 the goods are perfectly substituted and to increase the entropy they are independent.
The other two are dummies variables, where the first one indicates the production sector which seven dummies are identified, the other one indicates the Italian regions. In more details, the 21 dummies variables are: *Piemonte, Valle d'Aosta, Lombardia, Trentino Alto-Adige, Veneto, Friuli Venezia-Giulia, Liguria, Emilia Romagna, Toscana, Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia, Sardegna* and multi-region, the last one is assumed as a control variable. As before, regression is conducted on the dataset considering the two firms' networks, and then on the dataset related to the entire country.
It has been also considered as the dependent variable the logarithm of the distance ($\log(distance)$), the independent variables remained the same. Also, the same analysis is conducted on the same dataset and in the same way. For full knowledge, the followings are illustrated:
$\log(distance) = \alpha + \beta_1 n + \beta_2 H + \beta_3 sector + \beta_4 area + \epsilon$ &
$\log(distance) = \alpha + \beta_1 n + \beta_2 H + \beta_3 sector + \beta_4 region + \epsilon$
It is notable to point out that in some cases, networks are not included in the analysis: just in case the network average distance is equal to 0, and firms are in the same city, his $\log_3 0$ is equal to $-\infty$. It has no sense, so these statistical units are kept out of econometrics models.
The model and their variants aim to investigate if there is some relation between average distance and the nature of goods in stable networks to test if the [Di Dio et al., 2022] model is also sustained by empirical data. The econometric model tries to answer the following question: how does location choice between firms may influence R&D investment toward improvement in productivity and social welfare?
There is much literature evidence that firms prefer to maximise the positive externalities of spillovers. It is also evident that cooperation should be the focal point of each company; indeed, it dissuades the risk of opportunistic behaviour by rival companies.
All the results are exhibited in the next paragraph and the appendix.

3.5 Results

At this point, it is necessary to exhibit the results of our econometric models, taking the literature into account. The empirical analysis proposed by [Di Dio et al., 2022] investigates the networks composed of only two firms. This analysis is repeated and as well broadened to networks with n firms. In the tables, it is possible to see the main results:

With n=2

Dependent Variable: Average Distance						
Predictors	*[1]*	*[2]*	*[3]*	*[4]*	*[5]*	*[6]*
Intercept	47.53*** (4.66)	62.965*** (10.782)	325.517*** (5.803)	334.5592*** (8.4667)	334.8658*** (8.4962)	326.571*** (5.839)
Shannon Entropy	67.39*** (10.52)	43.761** (14.139)	8.871 (6.795)	-2.2286	-4.0638 (9.1461)	6.344 (6.918)
Normalised Entropy	-	-	-	-	-	-
Sector Dummies	-	Yes	-	Yes	Yes	-
Area Dummies	-	-	Yes	Yes	-	-
Region Dummies	-	-	-	-	Yes	Yes
Observations	2109	2109	2109	2109	2109	2109
F-Statistic	41.07*** (1; 2107 DF)	9.443*** (7; 2101 DF)	534.9*** (6; 2102 DF)	268.1*** (12; 2096 DF)	118.8*** (27; 2081 DF)	152.4*** (21; 2087)
R^2	0.01912	0.0305	0.6043	0.6055	0.6065	0.6053
Adj. R^2	0.01865	0.02727	0.6031	0.6032	0.6014	0.6013
AIC	27186.43	27221.81	25317.09	25430.45	26040.04	25746.57
BIC	27206.39	27227.7	25330.32	25369.6	25479.01	25746.53

*Note: Robust standard errors in parentheses.***p < .01; **p < .05; *p < .10.*

With n=2

Dependent Variable: Average Distance						
Predictors	*[7]*	*[8]*	*[9]*	*[10]*	*[11]*	*[12]*
Intercept	47.535*** (4.66)	62.965*** (10.782)	325.517*** (5.803)	334.5592*** (8.4667)	334.8658*** (8.4962)	326.571*** (5.839)
Shannon Entropy	-	-	-	-	-	-
Normalised Entropy	42.522*** (6.635)	27.610** (8.921)	5.597 (4.287)	-1.4061	-2.5640 (5.7706)	4.003 (4.365)
Sector Dummies	-	Yes	-	Yes	Yes	-
Area Dummies	-	-	Yes	Yes	-	-
Region Dummies	-	-	-	-	Yes	Yes
Observations	2109	2109	2109	2109	2109	2109
F-Statistic	41.07*** (1; 2107 DF)	9.443*** (7; 2101 DF)	534.9*** (6; 2102 DF)	268.1*** (12; 2096 DF)	118.8*** (27; 2081 DF)	152.4*** (21;2087 DF)
R^2	0.01912	0.0305	0.6043	0.6055	0.6065	0.6053
Adj. R^2	0.01865	0.02727	0.6031	0.6032	0.6014	0.6013
AIC	27186.43	27221.81	25317.09	25430.45	26040.04	25769.53
BIC	27206.39	27227.7	25330.32	25369.6	25479.01	25439.57

*Note: Robust standard errors in parentheses.***p < .01; **p < .05; *p < .10.*

FIGURE 3.9: Regression model's results with N=2, dividing per Shannon and normalised distance; dependent variable: average distance

As shown in the tables, there are many versions of the same model and some variables are not included in the analysis, there are two different tables; the first one (from models 1 to 6) takes into account the Shannon entropy parameter and the second one (from models 7 to 12) considers the normalised entropy; these model wants to check the empirical evidence of [Di Dio et al., 2022] research.

It is possible to note that all the models are statistically significant at the 95% level of significance. Both territorial variables are statistically significant at the 95% level of significance and negatively affect the distance standing for the firms in the same networks is about 300 km closer than the multi-regional networks. In most cases, the sector variables are not statistically significant at the 95% level of significance, except for the agriculture sector, which negatively affects the distance. Entropy is not always statically significant at the 95% level of significance, so there is not sufficient evidence to confirm that is a correspondence between the produced good and the network average distance.

Different criteria try to identify the best models, considering the R^2, the best models are the fifth and the eleventh, as the $\bar{R}^2$ best models are the fourth and tenth considering the AIC and BIC criteria, best models are the third and the ninth. Giving up some accuracy, but gaining some information: the fourth and tenth models are preferable. Considering these last two models, entropy is negative, which means the goods are perfectly substituted. They will tend to maximise the distance to relax competition and minimise positive externalities in favour of the rival firm [Di Dio et al., 2022], instead of a two-firm network, it is not possible to confirm the weak empirical evidence of [Di Dio et al., 2022].

The same analysis is repeated considering as the dependent variable the logarithm of distance to verify if the different parameters produce a percentage change of distance, given a unitary increase in the kind of goods produced. In the case of dummy variables, the goal is to measure the percentage difference of the average distance between sectors and multi-sector variables and between the region and multi-region variables.

From 1 to 6 models, the Shannon entropy variable is applied and from 7 to 12 models, normalised entropy is.

It is remarkable to point out that when the average distance is equal to 0 and his logarithm is equal to $-\infty$, 957 observations are eliminated.

With n=2

Dependent Variable: Logarithm of Average Distance

	[1]	[2]	[3]	[4]	[5]	[6]
Predictors						
Intercept	3.54240*** (0.06547)	3.6036*** (0.13413)	5.41639*** (0.05976)	5.42741*** (0.09526)	5.44773*** (0.094714)	5.4437*** (0.05955)
Shannon Entropy	0.70484*** (0.13814)	0.53664** (0.17923)	0.20946 (0.08693)	0.14883 (0.11322)	0.091305 (0.113397)	0.144 (0.08775)
Normalised Entropy	-	-	-	-	-	-
Sector Dummies	-	Yes	-	Yes	Yes	-
Area Dummies	-	-	Yes	Yes	-	-
Region Dummies	-	-	-	-	Yes	Yes
Observations	2109	2109	2109	2109	2109	2109
F-Statistic	26.03*** (1; 1150 DF)	11.8*** (7; 1144 DF)	319.8*** (6; 1145)	163.7*** (12; 1139 DF)	75.67*** (27; 1124 DF)	95.4*** (21; 1130 DF)
R^2	0.02214	0.06733	0.6263	0.633	0.6451	0.6394
Adj. R^2	0.02129	0.06162	0.6243	0.6292	0.6366	0.6327
AIC	4152.986	4158.473	3089.963	3188.918	3765.396	3483.84
BIC	4171.134	4158.917	3098.357	3119.607	3186.824	3162.973

*Note: Robust standard errors in parentheses.***p < .01; **p < .05; *p < .10.*
(957 eliminated observation because of $log_3(0)$ is equal to -Inf)

With n=2

Dependent Variable: Logarithm of Average Distance

	[7]	[8]	[9]	[10]	[11]	[12]
Predictors						
Intercept	3.5424*** (0.06547)	3.0603*** (0.13413)	5.41639*** (0.05976)	5.42741*** (0.09526)	5.44477*** (0.094714)	5.4437*** (0.05955)
Shannon Entropy	-	-	-	-	-	-
Normalised Entropy	0.4447*** (0.08716)	0.33858** (0.11308)	0.013215* (0.05484)	0.0939 (0.07143)	0.05744 (0.071546)	0.09085 (0.05537)
Sector Dummies	-	Yes	-	Yes	Yes	-
Area Dummies	-	-	Yes	Yes	-	-
Region Dummies	-	-	-	-	Yes	Yes
Observations	2109	2109	2109	2109	2109	2109
F-Statistic	26.03*** (1; 1150 DF)	11.8*** (7; 1144 DF)	319.8*** (6; 1145)	163.7*** (12; 1139 DF)	75.67*** (27; 1124 DF)	95.4*** (21; 1130 DF)
R^2	0.02214	0.06733	0.6263	0.633	0.6451	0.6394
Adj. R^2	0.02129	0.06162	0.6243	0.6292	0.6366	0.6327
AIC	4152.986	4158.473	3089.963	3188.918	3765.396	3483.84
BIC	4171.134	4158.917	3098.357	3119.607	3186.824	3162.973

*Note: Robust standard errors in parentheses.***p < .01; **p < .05; *p < .10.*
(957 eliminated observation because of $log_3(0)$ is equal to -Inf)

Figure 3.10: Regression model's results with N=2, dividing per Shannon and normalised distance; dependent variable: logarithm of average distance

The log-linear regressions are all statistically significant at the 95% level of significance. Both territorial variables are statistically significant at 95 %, and the firms in regional networks are closer (about 200-300%) than firms in multi-region. Entropy is not always statically significant at the 95% level of significance, so there is not sufficient evidence to confirm that exists a correspondence between the produced good and the network's average distance.

For defining the best model, the same criteria were applied. The fifth and eleventh models have a higher R^2 and $\bar{R}^2$; according to BIC and AIC criteria, the third and ninth models are better (lower values). Giving up some accuracy but gaining some information: the better models are the fifth and the ninth. Considering these two models' entropy is positive: it produces a percentage change in distance respectively by 9.13 and 1.13 given an increment of both entropy by 1 unit.

In this case, the empirical evidence cannot confirm [Di Dio et al., 2022] one. Indeed, in the case of substituted goods, companies tend to maximise positive externalities in favour of the rival firm [Di Dio et al., 2022].

In both proposed regression models, there is no empirical evidence to reject the other result from the model that affirm that two firms tend to form stable cooperative link irrespective of the nature of the goods they produce [Di Dio et al., 2022].

The empirical research has broadened considering all the Italian business networks. As above, from 1 to 6 models Shannon entropy is considered instead of from 7 to 12 normalised entropy.

Dependent Variable: Average Distance

	[1]	[2]	[3]	[4]	[5]	[6]
Predictors						
Intercept	53.2524*** (2.4505)	71.9097*** (4.9969)	307.1818*** (4.5832)	293.7525*** (5.4501)	293.7026*** (5.4438)	307.4646*** (4.5699)
Number of Firms	0.2463 (0.159)	0.6112*** (0.1616)	0.7589*** (0.1276)	0.8486*** (0.13)	0.8512*** (0.1303)	0.7873*** (0.1281)
Shannon Entropy	28.5282*** (3.1786)	16.6658*** (3.3675)	32.5991*** (2.5503)	29.0129*** (2.1791)	29.0972*** (2.7262)	31.87*** (2.5619)
Normalised Entropy	-	-	-	-	-	-
Sector Dummies	-	Yes	-	Yes	Yes	-
Area Dummies	-	-	Yes	Yes	-	-
Region Dummies	-	-	-	-	Yes	Yes
Observations	6749	6749	6749	6749	6749	6749
F-Statistic	46.09*** (2; 6746 DF)	28.49*** (8; 6740 DF)	562.8*** (7; 6741 DF)	312.5*** (13; 6735 DF)	147.6*** (28; 6720 DF)	183.1*** (22; 6726 DF)
R^2	0.01348	0.03271	0.3689	0.3672	0.3808	0.3745
Adj. R^2	0.01319	0.03156	0.3682	0.375	0.3783	0.3725
AIC	84606.58	84545.71	81647.13	81699.85	82294.67	82051.17
BIC	84633.85	84553.89	81663.49	81637.1	81719.18	81734.79

*Note: Robust standard errors in parentheses.***p < .01; **p < .05; *p < .10.*

Dependent Variable: Average Distance

	[7]	[8]	[9]	[10]	[11]	[12]
Predictors						
Intercept	49.8536*** (2.6686)	65.7173*** (6.0819)	310.8943*** (4.8808)	296.4681*** (6.2823)	296.916*** (6.279)	311.7759*** (4.8702)
Number of firms	0.6516*** (0.1562)	0.8213*** (0.1563)	1.1392*** (0.1266)	1.2031*** (0.1267)	1.207*** (0.127)	1.1605*** (0.1269)
Shannon Entropy	-	-	-	-	-	-
Normalised Entropy	36.205*** (3.8873)	19.5435*** (4.4266)	18.8890*** (3.1581)	15.7878*** (3.6641)	15.325*** (3.674)	17.3575*** (3.1755)
Sector Dummies	-	Yes	-	Yes	Yes	-
Area Dummies	-	-	Yes	Yes	-	-
Region Dummies	-	-	-	-	Yes	Yes
Observations	6749	6749	6749	6749	6749	6749
F-Statistic	49.19*** (2; 6746 DF)	27.74*** (8; 6740 DF)	534.6*** (7; 6741)	300.9*** (13; 6735 DF)	142.2*** (28; 6720 DF)	174.2*** (22; 6276 DF)
R^2	0.01437	0.03187	0.357	0.3674	0.372	0.363
Adj. R^2	0.01408	0.03073	0.3563	0.3662	0.3694	0.3609
AIC	84600.47	84551.56	81773.05	81794.4	82390.67	82174.79
BIC	84627.74	84559.73	81789.41	81731.66	81815.19	81858.4

*Note: Robust standard errors in parentheses.***p < .01; **p < .05; *p < .10.*

FIGURE 3.11: Regression model's results with all Italian business networks, dividing per Shannon and normalised distance; dependent variable: average distance

The networks are 6749. All the statistical models are statistically significant at the 95% level of significance. A new independent variable, the firms' average number per network, is included in the econometric model. The number of firms, type of goods, territorial variables and intercept is all statistically significant at the 95% level of significance. Both entropies are positive: if the maximum dispersion of firms holds on the condition that goods are independent of each other, the statement is opposite to [Shimizu, 2002] results. The territorial variables negatively affect the distance: it is obvious that networks in the same region are 300 km closer than the multi-regional networks.

According to the criteria that explain the percentage of the variance explained by the model ($R^2 \quad and \quad \bar{R^2}$), the fifth and eleventh models are better; they explain about 37% of variance. Rather than AIC and BIC criteria, the third and ninth models are the better ones, but they do not take into consideration the sector variables, which is why we prefer the earlier models.

As above, the analysis is repeated considering as the dependent variable the logarithm of distance. The results are shown with the same division.

It is also crucial for a correct data interpretation that 1743 observations cannot be considered in the regression because their average distance is equal to 0, and its logarithm is equal to $-\infty$ and has no sense in economics. Intercept, number of firms, territorial, and entropy variables are statistically significant at 95% level of significance, that means when entropy or the number of firms increases by 1 unit, average distance increase respectively by 36.08% (Shannon Entropy) or 28.62% (Normalised entropy) and about 1%.

Examining all the models, the fifth and the eleventh ones are better considering R^2 and $\bar{R^2}$, instead regarding AIC and BIC criteria, the fourth and tenth are the better ones, as before, they are not preferred to the others because they ignore the sector variable. It is preferable to give up on accuracy, but gain information.

The results are shown in the following tables.

Dependent Variable: Logarithm of Average Distance

	[1]	[2]	[3]	[4]	[5]	[6]
Predictors						
Intercept	3.14812*** (0.03244)	3.66168*** (0.06439)	5.28768*** (0.05488)	5.13561*** (0.06831)	5.14422*** (0.06784)	5.29675*** (0.05431)
Number of Firms	0.00115 (0.001756)	0.00625*** (0.00177)	0.00633*** (0.00154)	0.00872*** (0.00156)	0.00854*** (0.00156)	0.00664*** (0.00154)
Shannon Entropy	0.3517*** (0.03892)	0.19467*** (0.04047)	0.45239*** (0.03413)	0.36802*** (0.03592)	0.36084*** (0.03585)	0.432621*** (0.03406)
Normalised Entropy	-	-	-	-	-	-
Sector Dummies	-	Yes	-	Yes	Yes	-
Area Dummies	-	-	Yes	Yes	-	-
Region Dummies	-	-	-	-	Yes	Yes
Observations	6749	6749	6749	6749	6749	6749
F-Statistic	43.49*** (2; 5003 DF)	40.1*** (8; 4997 DF)	243.2 (7; 4998 DF)	147.3*** (13; 4992)	73.02*** (28; 4977 DF)	84.93*** (22; 4983 DF)
R^2	0.01709	0.0633	0.2541	0.2772	0.2912	0.2727
Adj. R^2	0.0167	0.05882	0.253	0.2753	0.2872	0.2695
AIC	17487.62	17334.42	16161.57	16135.89	16683	16499.84
BIC	17513.69	17339.6	16175.24	16068.66	16098.55	16176.28

*Note: Robust standard errors in parentheses.***p < .01; **p < .05; *p < .10.*
(1743 eliminated observation because of $log_3(0)$ is equal to -Inf)

Dependent Variable: Logarithm of Average Distance

	[7]	[8]	[9]	[10]	[11]	[12]
Predictors						
Intercept	3.33975*** (0.03561)	3.5186*** (0.07818)	5.26444***(0.06017)	5.07651*** (0.0809)	5.09615*** (0.0804)	5.28574*** (0.05962)
Number of firms	0.00611*** (0.00174)	0.00861*** (0.00172)	0.01122*** (0.00154)	0.012711*** (0.00153)	0.01243*** (0.00153)	0.0113*** (0.00154)
Shannon Entropy	-	-	-	-	-	-
Normalised Entropy	0.537742*** (0.05135)	0.31381*** (0.05818)	0.38491*** (0.0456)	0.30346***(0.05145)	0.28662*** (0.05139)	0.35004*** (0.04557)
Sector Dummies	-	Yes	-	Yes	Yes	-
Area Dummies	-	-	Yes	Yes	-	-
Region Dummies	-	-	-	-	Yes	Yes
Observations	6749	6749	6749	6749	6749	6749
F-Statistic	57.51*** (2; 5003 DF)	40.89*** (8; 4997 DF)	223.9*** (7; 4998 DF)	139.9*** (13; 4992 DF)	69.59***(28; 4977 DF)	78.73*** (22; 4983 DF)
R^2	0.02247	0.06144	0.2387	0.2671	0.2813	0.2579
Adj. R^2	0.02208	0.05994	0.2376	0.2652	0.2772	0.2547
AIC	17460.12	17328.48	16263.67	16205.36	16752.67	16600.43
BIC	17486.2	17333.66	16277.34	16138.14	16168.22	16276.88

*Note: Robust standard errors in parentheses.***p < .01; **p < .05; *p < .10.*
(1743 eliminated observation because of $log_3(0)$ is equal to -Inf)

FIGURE 3.12: Regression model's results with all Italian business networks, dividing per Shannon and normalised distance; dependent variable: average distance

In conclusion, it is possible to see that these econometric models are very effective, especially the fifth one applied to all Italian business networks, which considers more territorial and sector variables.
At this point of the thesis, the results are contextualised with the literature. There is enough evidence to reject the assumption of [Di Dio et al., 2022], opposite results appeared, the nature of goods produced affects the distance between firms of the same networks. It tends to be broader when companies produce independent goods; instead of the first hypothesis of the model [Di Dio et al., 2022] being confirmed, networks arise independently of the nature of goods by participating firms.
The empirical results suggest that when firms tend to agglomerate, they produce substituted goods; they want to take advantage of positive externalities that are typical of industrial districts. It is also worthy of mention that restricted context permits greater stability of collusive agreement to maintain a high price.
The econometric model confirms the [Antonietti et al., 2019] results, in which entropy is used to measure the economic complexity, that is the variety of industries and kinds of produced goods. The greater the economic complexity, the more the development of a region is.
In the next chapter, the conclusion and future work are proposed.

Chapter 4

Conclusion and Future Works

As investigated by [Marshall, 2009], industrial districts produce positive externalities and knowledge spillover. They are very important for the Italian economy [Zazzaro, 2010]; indeed; they aim to improve productivity, as mentioned in [Brandolini and Bugamelli, 2009]. They contribute significantly to Italian growth and development [Cassa Depositi e Prestiti, 2018]. TFP is stagnant, but any form of collaboration between enterprises may be an instrument to improve themselves: increasing international trade, and productivity [Hulten, 2001]. In particular, business networks can be seen as an extension of industrial districts, though they cannot be used as synonyms [Garofalo and Pugliesi, 2014].
Also [Fabrizi et al., 2022] analyse their importance for policymakers, which should take into account inter-firms cooperation. A more collaborative environment between firms may reduce the gap between the north and south of Italy [Pastore et al., 2020].
As seen, business networks are a very strategic instrument for Italian development; some economical models should investigate the impact of business networks on Italian GDP. The phenomenon has assumed great importance, and it takes on macro-economic relevance: a policy considering all the complexity of the Italian production system should be proposed.
As mentioned in [Costa et al., 2017], attendance in a business network contract increments the revenue for the same enterprise and produces positive externalities for the area where the company is based.
The results of this empirical analysis suggest that when firms tend to agglomerate, then they produce substitute goods, despite [Di Dio et al., 2022]. For this reason, it is possible to suggest that similar firms are closer to taking maximum advantage of network

externalities. A greater concentration of positive externalities is typical of industrial districts, in addition, similar activities in a restricted context permit greater stability of collusive agreement to maintain a high price.
It should be thought-inspiring to investigate how business networks can contribute to the business' revenues and try to develop a model that can investigate how human relations may influence productivity.
Future models may also include other predictors such as the firm's revenue or other profitability indexes.
It is also of interest to expand and revise the model of [Di Dio et al., 2022] considering not only two-firm networks but collaborations among N firms. For example, it could be investigated whether the dimension of firms or networks influences the kind of goods produced. As mentioned above, independent goods tend to increase the distance between network firms much more than in the case of substitute or complementary goods, i. e. agglomeration is greater. As evidenced by the low portion variance explained by the model, other variables may be added.
It is known that industrial districts were born before business network contracts, and it should be very interesting to explain how business networks impact industrial districts.
In conclusion, it is of great relevance to affirm the importance of business network contracts, and how they are a valid instrument for competitiveness and efficiency of productivity, if which policymakers should be more conscious.

Appendix A

Ateco Code Classification

APPENDICE

Classificazione delle imprese in base ai macro-ambiti di attività

MACRO-AMBITO	CODICE ATECO - 2 DIGIT
Agroalimentare	A1, A2, A3, C10, C11
Meccanica	C24, C25, C27, C28, C29, C30, C33
Costruzioni	C23, F41, F42, F43, L68
Commercio	G45, G46, G47
Sistema moda e arredo	C13, C14, C15, C16, C31
Turismo	I55, I56, N79,
Trasporti e logistica	H49, H50, H51, H52, H53
Servizi operativi	N77, N78, N80, N81, N82
Servizi professionali	M69, M70, M71, M72, M73, M74
Servizi tecnologici, di informazione e comunicazione	J61, J60, J59, J58, J62, J63
Sanità	Q86, Q87, Q88
Servizi finanziari	K64, K65, K66
Attività immobiliari	L68
Attività di formazione, culturali e di intrattenimento	R90, R91, R92, R93, P85
Servizi per la persona	S96, S95
Altre attività manifatturiere	C12, C17, C18, C19, C20, C21, C22, C24, C26, C32
Utilities e servizi ambientali	D35, E36, E37, E38, E39
Altro	B08, B09, M75, S94, NS, O84

Elaborazione RetImpresa su Codici Ateco 2007

FIGURE A.1: Ateco code classification [Pitingaro and Corsini, 2021, RetImpresa, 2019a, RetImpresa, 2020]

Appendix B

Regression Results

TABLE B.1: Regression Models with N=2, Dep. Var.: Average Distance (Shannon Entropy) - A

	Dependent variable:		
	Average distance		
	(1)	(2)	(3)
Shannon_entropia	67.395^{***}	43.761^{***}	8.871
	(10.517)	(14.139)	(6.795)
d_agr		-49.849^{***}	
		(14.007)	
d_altro		-2.054	
		(54.699)	
d_comm		-25.200	
		(17.492)	
d_ind		-7.826	
		(10.640)	
d_serv		6.591	
		(10.226)	
d_tur		-3.866	
		(34.605)	
nord_ovest			-315.824^{***}
			(7.129)
nord_est			-315.227^{***}
			(6.815)
centro			-315.626^{***}
			(7.087)
sud			-311.010^{***}
			(7.082)
isole			-311.783^{***}
			(11.134)
Constant	47.535^{***}	62.965^{***}	325.517^{***}
	(4.660)	(10.782)	(5.803)
Observations	2,109	2,109	2,109
R^2	0.019	0.031	0.604
Adjusted R^2	0.019	0.027	0.603
Residual Std. Error	152.347 (df = 2107)	151.677 (df = 2101)	96.883 (df = 2102)
F Statistic	41.066^{***} (df = 1; 2107)	9.443^{***} (df = 7; 2101)	534.932^{***} (df = 6; 2102)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.2: Regression Models with N=2, Dep. Var: Average Distance (Shannon Entropy) - B

	Dependent variable:		
	Average Distance		
	(4)	(5)	(6)
Shannon_entropia	−2.229	−4.064	6.344
	(9.071)	(9.146)	(6.918)
d_agr	-18.388^{**}	-17.622^{*}	
	(9.001)	(9.125)	
d_altro	−0.308	0.927	
	(35.019)	(35.179)	
d_comm	-19.425^{*}	-19.290^{*}	
	(11.182)	(11.254)	
d_ind	−3.196	−2.315	
	(6.837)	(6.903)	
d_serv	−7.142	−6.434	
	(6.539)	(6.603)	
d_tur	−3.384	−3.973	
	(22.164)	(22.252)	
nord_ovest	-315.184^{***}		
	(7.162)		
nord_est	-314.303^{***}		
	(6.867)		
centro	-314.621^{***}		
	(7.132)		
sud	-310.557^{***}		
	(7.101)		
isole	-310.875^{***}		
	(11.177)		
d_piemonte		-317.794^{***}	-318.897^{***}
		(9.915)	(9.846)
d_valle_d_aosta		-325.809^{***}	-322.921^{***}
		(32.824)	(32.763)
d_lombardia		-312.448^{***}	-312.961^{***}
		(8.525)	(8.503)
d_trentino		-316.821^{***}	-317.713^{***}
		(14.126)	(14.103)
d_veneto		-314.505^{***}	-315.237^{***}
		(8.418)	(8.351)
d_friuli		-310.751^{***}	-313.214^{***}
		(11.533)	(11.424)
d_liguria		-321.751^{***}	-322.358^{***}
		(17.441)	(17.436)
d_emilia		-317.246^{***}	-317.308^{***}
		(10.613)	(10.600)
d_toscana		-313.633^{***}	-315.405^{***}
		(10.307)	(10.189)
d_umbria		-318.400^{***}	-319.528^{***}
		(18.871)	(18.772)
d_marche		-318.695^{***}	-317.033^{***}
		(14.168)	(14.103)
d_lazio		-314.375^{***}	-315.616^{***}
		(8.711)	(8.678)
d_abruzzo		-319.682^{***}	-320.795^{***}
		(11.753)	(11.680)
d_molise		-321.593^{***}	-319.315^{***}
		(48.871)	(48.848)
d_campania		-316.435^{***}	-316.852^{***}
		(9.885)	(9.873)
d_puglia		-295.981^{***}	-295.810^{***}
		(10.083)	(10.048)
d_basilicata		-317.893^{***}	-320.557^{***}
		(28.556)	(28.495)
d_calabria		-317.950^{***}	-319.666^{***}
		(19.833)	(19.782)
d_sicilia		-314.620^{***}	-315.285^{***}
		(13.565)	(13.551)
d_sardegna		-305.193^{***}	-306.676^{***}
		(17.065)	(16.971)
Constant	334.559^{***}	334.866^{***}	326.571^{***}
	(8.467)	(8.496)	(5.839)
Observations	2,109	2,109	2,109
R^2	0.606	0.607	0.605
Adjusted R^2	0.603	0.601	0.601
Residual Std. Error	96.868 (df = 2096)	97.092 (df = 2081)	97.102 (df = 2087)
F Statistic	268.095^{***} (df = 12; 2096)	118.802^{***} (df = 27; 2081)	152.411^{***} (df = 21; 2087)
Note:			*p<0.1; **p<0.05; ***p<0.01

TABLE B.3: Regression Models with N=2, Dep. Var.: Average distance (Normalised Entropy) - A

	Dependent variable:		
	Average distance		
	(7)	(8)	(9)
H_norm	42.522^{***}	27.610^{***}	5.597
	(6.635)	(8.921)	(4.287)
d_agr		-49.849^{***}	
		(14.007)	
d_altro		-2.054	
		(54.699)	
d_comm		-25.200	
		(17.492)	
d_ind		-7.826	
		(10.640)	
d_serv		6.591	
		(10.226)	
d_tur		-3.866	
		(34.605)	
nord_ovest			-315.824^{***}
			(7.129)
nord_est			-315.227^{***}
			(6.815)
centro			-315.626^{***}
			(7.087)
sud			-311.010^{***}
			(7.082)
isole			-311.783^{***}
			(11.134)
Constant	47.535^{***}	62.965^{***}	325.517^{***}
	(4.660)	(10.782)	(5.803)
Observations	2,109	2,109	2,109
R^2	0.019	0.031	0.604
Adjusted R^2	0.019	0.027	0.603
Residual Std. Error	152.347 (df = 2107)	151.677 (df = 2101)	96.883 (df = 2102)
F Statistic	41.066^{***} (df = 1; 2107)	9.443^{***} (df = 7; 2101)	534.932^{***} (df = 6; 2102)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.4: Regression models with N=2, Dep. Var. Average distance (Normalised Entropy) - B

	Dependent variable:		
	Average distance		
	(10)	(11)	(12)
Shannon_entropia	-2.229		
	(9.071)		
H_norm		-2.564	4.003
		(5.771)	(4.365)
d_agr	-18.388^{**}	-17.622^{*}	
	(9.001)	(9.125)	
d_altro	-0.308	0.927	
	(35.019)	(35.179)	
d_comm	-19.425^{*}	-19.290^{*}	
	(11.182)	(11.254)	
d_ind	-3.196	-2.315	
	(6.837)	(6.903)	
d_serv	-7.142	-6.434	
	(6.539)	(6.603)	
d_tur	-3.384	-3.973	
	(22.164)	(22.252)	
nord_ovest	-315.184^{***}		
	(7.162)		
nord_est	-314.303^{***}		
	(6.867)		
centro	-314.621^{***}		
	(7.132)		
sud	-310.557^{***}		
	(7.101)		
isole	-310.875^{***}		
	(11.177)		
d_piemonte		-317.794^{***}	-318.897^{***}
		(9.915)	(9.846)
d_valle_d_aosta		-325.809^{***}	-322.921^{***}
		(32.824)	(32.763)
d_lombardia		-312.448^{***}	-312.961^{***}
		(8.525)	(8.503)
d_trentino		-316.821^{***}	-317.713^{***}
		(14.126)	(14.103)
d_veneto		-314.505^{***}	-315.237^{***}
		(8.418)	(8.351)
d_friuli		-310.751^{***}	-313.214^{***}
		(11.533)	(11.424)
d_liguria		-321.751^{***}	-322.358^{***}
		(17.441)	(17.436)
d_emilia		-317.246^{***}	-317.308^{***}
		(10.613)	(10.600)
d_toscana		-313.633^{***}	-315.405^{***}
		(10.307)	(10.189)
d_umbria		-318.400^{***}	-319.528^{***}
		(18.871)	(18.772)
d_marche		-318.695^{***}	-317.033^{***}
		(14.168)	(14.103)
d_lazio		-314.375^{***}	-315.616^{***}
		(8.711)	(8.678)
d_abruzzo		-319.682^{***}	-320.795^{***}
		(11.753)	(11.680)
d_molise		-321.593^{***}	-319.315^{***}
		(48.871)	(48.848)
d_campania		-316.435^{***}	-316.852^{***}
		(9.885)	(9.873)
d_puglia		-295.981^{***}	-295.810^{***}
		(10.083)	(10.048)
d_basilicata		-317.893^{***}	-320.557^{***}
		(28.556)	(28.495)
d_calabria		-317.950^{***}	-319.666^{***}
		(19.833)	(19.782)
d_sicilia		-314.620^{***}	-315.285^{***}
		(13.565)	(13.551)
d_sardegna		-305.193^{***}	-306.676^{***}
		(17.065)	(16.971)
Constant	334.559^{***}	334.866^{***}	326.571^{***}
	(8.467)	(8.496)	(5.839)
Observations	2,109	2,109	2,109
R^2	0.606	0.607	0.605
Adjusted R^2	0.603	0.601	0.601
Residual Std. Error	96.868 (df = 2096)	97.092 (df = 2081)	97.102 (df = 2087)
F Statistic	268.095^{***} (df = 12; 2096)	118.802^{***} (df = 27; 2081)	152.411^{***} (df = 21; 2087)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.5: Regression Models with N=2, Dep. Var.: log(*Average Distance*) (Shannon Entropy) - A

	Dependent variable:		
	log(*Average Distance*)		
	(1)	(2)	(3)
Shannon_entropia	0.705***	0.537***	0.149
	(0.138)	(0.179)	(0.113)
d_agr		−0.653***	−0.226*
		(0.188)	(0.120)
d_altro		0.622	0.063
		(1.024)	(0.647)
d_comm		0.092	−0.011
		(0.243)	(0.153)
d_ind		−0.194	−0.137*
		(0.128)	(0.082)
d_serv		0.378***	0.132*
		(0.126)	(0.079)
d_tur		1.427**	0.661
		(0.653)	(0.411)
nord_ovest			−2.581***
			(0.083)
nord_est			−2.462***
			(0.079)
centro			−2.267***
			(0.088)
sud			−2.417***
			(0.082)
isole			−2.530***
			(0.140)
Constant	3.542***	3.604***	5.427***
	(0.065)	(0.134)	(0.095)
Observations	1,152	1,152	1,152
R^2	0.022	0.067	0.633
Adjusted R^2	0.021	0.062	0.629
Residual Std. Error	1.467 (df = 1150)	1.436 (df = 1144)	0.903 (df = 1139)
F Statistic	26.032*** (df = 1; 1150)	11.798*** (df = 7; 1144)	163.734*** (df = 12; 1139)
Note:			*p<0.1; **p<0.05; ***p<0.01

TABLE B.6: Regression Model with N=2, Dep. Var.: log(*Average Distance*) (Shannon Entropy) - B

	Dependent variable:		
	log(*Average Distance*)		
	(4)	(5)	(6)
Shannon_entropia	0.209**	0.091	0.144
	(0.087)	(0.113)	(0.088)
nord_ovest	−2.641***		
	(0.083)		
nord_est	−2.529***		
	(0.078)		
centro	−2.307***		
	(0.088)		
sud	−2.427***		
	(0.082)		
isole	−2.527***		
	(0.140)		
d_agr		−0.222*	
		(0.121)	
d_altro		0.089	
		(0.644)	
d_comm		−0.007	
		(0.153)	
d_ind		−0.103	
		(0.082)	
d_serv		0.130	
		(0.079)	
d_tur		0.651	
		(0.408)	
d_piemonte		−2.641***	−2.713***
		(0.134)	(0.133)
d_valle_d_aosta		−3.477***	−3.544***
		(0.403)	(0.405)
d_lombardia		−2.511***	−2.563***
		(0.097)	(0.096)
d_trentino		−2.409***	−2.427***
		(0.201)	(0.202)
d_veneto		−2.534***	−2.598***
		(0.101)	(0.101)
d_friuli		−2.360***	−2.461***
		(0.141)	(0.139)
d_liguria		−2.800***	−2.835***
		(0.287)	(0.288)
d_emilia		−2.486***	−2.530***
		(0.140)	(0.140)
d_toscana		−2.377***	−2.425***
		(0.136)	(0.136)
d_umbria		−2.483***	−2.538***
		(0.290)	(0.288)
d_marche		−2.493***	−2.535***
		(0.183)	(0.183)
d_lazio		−2.087***	−2.114***
		(0.119)	(0.120)
d_abruzzo		−2.648***	−2.708***
		(0.168)	(0.168)
d_molise		−2.656***	−2.642***
		(0.635)	(0.638)
d_campania		−2.725***	−2.736***
		(0.124)	(0.124)
d_puglia		−2.014***	−2.010***
		(0.116)	(0.116)
d_basilicata		−2.510***	−2.618***
		(0.405)	(0.405)
d_calabria		−2.651***	−2.581***
		(0.304)	(0.304)
d_sicilia		−2.450***	−2.417***
		(0.189)	(0.190)
d_sardegna		−2.615***	−2.647***
		(0.191)	(0.189)
Constant	5.416***	5.445***	5.444***
	(0.060)	(0.095)	(0.060)
Observations	1,152	1,152	1,152
R^2	0.626	0.645	0.639
Adjusted R^2	0.624	0.637	0.633
Residual Std. Error	0.909 (df = 1145)	0.894 (df = 1124)	0.899 (df = 1130)
F Statistic	319.779*** (df = 6; 1145)	75.670*** (df = 27; 1124)	95.401*** (df = 21; 1130)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.7: Regression Model with N=2, Dep. Var.: log(*Average Distance*) (Normalised Entropy) - A

	Dependent variable:		
	log(*Average Distance*)		
	(7)	(8)	(9)
H_norm	0.445^{***}	0.339^{***}	0.132^{**}
	(0.087)	(0.113)	(0.055)
d_agr		-0.653^{***}	
		(0.188)	
d_altro		0.622	
		(1.024)	
d_comm		0.092	
		(0.243)	
d_ind		-0.194	
		(0.128)	
d_serv		0.378^{***}	
		(0.126)	
d_tur		1.427^{**}	
		(0.653)	
nord_ovest			-2.641^{***}
			(0.083)
nord_est			-2.529^{***}
			(0.078)
centro			-2.307^{***}
			(0.088)
sud			-2.427^{***}
			(0.082)
isole			-2.527^{***}
			(0.140)
Constant	3.542^{***}	3.604^{***}	5.416^{***}
	(0.065)	(0.134)	(0.060)
Observations	1,152	1,152	1,152
R^2	0.022	0.067	0.626
Adjusted R^2	0.021	0.062	0.624
Residual Std. Error	1.467 (df = 1150)	1.436 (df = 1144)	0.909 (df = 1145)
F Statistic	26.032^{***} (df = 1; 1150)	11.798^{***} (df = 7; 1144)	319.779^{***} (df = 6; 1145)
Note:			*p<0.1; **p<0.05; ***p<0.01

TABLE B.8: Regression Model with N=2, Dep. Var.:log(*Average Distance*) (Normalised Entropy) - B

	Dependent variable: log(*Average Distance*) (10)	(11)	(12)
Shannon_entropia	0.209** (0.087)		
nord_ovest	−2.641*** (0.083)		
nord_est	−2.529*** (0.078)		
centro	−2.307*** (0.088)		
sud	−2.427*** (0.082)		
isole	−2.527*** (0.140)		
H_norm		0.057 (0.072)	0.091 (0.055)
d_agr		−0.222* (0.121)	
d_altro		0.089 (0.644)	
d_comm		−0.007 (0.153)	
d_ind		−0.103 (0.082)	
d_serv		0.130 (0.079)	
d_tur		0.651 (0.408)	
d_piemonte		−2.641*** (0.134)	−2.713*** (0.133)
d_valle_d_aosta		−3.477*** (0.403)	−3.544*** (0.405)
d_lombardia		−2.511*** (0.097)	−2.563*** (0.096)
d_trentino		−2.409*** (0.201)	−2.427*** (0.202)
d_veneto		−2.534*** (0.101)	−2.598*** (0.101)
d_friuli		−2.360*** (0.141)	−2.461*** (0.139)
d_liguria		−2.800*** (0.287)	−2.835*** (0.288)
d_emilia		−2.486*** (0.140)	−2.530*** (0.140)
d_toscana		−2.377*** (0.136)	−2.425*** (0.136)
d_umbria		−2.483*** (0.290)	−2.538*** (0.288)
d_marche		−2.493*** (0.183)	−2.535*** (0.183)
d_lazio		−2.087*** (0.119)	−2.114*** (0.120)
d_abruzzo		−2.648*** (0.168)	−2.708*** (0.168)
d_molise		−2.656*** (0.635)	−2.642*** (0.638)
d_campania		−2.725*** (0.124)	−2.736*** (0.124)
d_puglia		−2.014*** (0.116)	−2.010*** (0.116)
d_basilicata		−2.510*** (0.405)	−2.618*** (0.405)
d_calabria		−2.651*** (0.304)	−2.581*** (0.304)
d_sicilia		−2.450*** (0.189)	−2.417*** (0.190)
d_sardegna		−2.615*** (0.191)	−2.647*** (0.189)
Constant	5.416*** (0.060)	5.445*** (0.095)	5.444*** (0.060)
Observations	1,152	1,152	1,152
R^2	0.626	0.645	0.639
Adjusted R^2	0.624	0.637	0.633
Residual Std. Error	0.909 (df = 1145)	0.894 (df = 1124)	0.899 (df = 1130)
F Statistic	319.779*** (df = 6; 1145)	75.670*** (df = 27; 1124)	95.401*** (df = 21; 1130)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.9: Regression results, Dep. Var. Average Distance (Shannon Entropy) - A

	Dependent variable:		
	Average Distance		
	(1)	(2)	(3)
n	0.246	0.611***	0.849***
	(0.159)	(0.162)	(0.130)
Shannon_entropia	28.528***	16.666***	29.013***
	(3.179)	(3.367)	(2.719)
d_agr		−49.794***	−0.872
		(6.161)	(5.021)
d_altro		−9.448	22.568
		(22.435)	(18.078)
d_comm		−32.274***	3.693
		(7.409)	(5.984)
d_ind		−14.256***	17.242***
		(5.223)	(4.240)
d_serv		3.645	26.821***
		(5.131)	(4.144)
d_tur		−23.856**	8.999
		(11.010)	(8.873)
nord_ovest			−275.434***
			(5.242)
nord_est			−289.357***
			(5.123)
centro			−288.431***
			(5.145)
sud			−272.027***
			(5.272)
isole			−255.215***
			(7.105)
Constant	53.252***	71.910***	293.753***
	(2.451)	(4.997)	(5.450)
Observations	6,749	6,749	6,749
R^2	0.013	0.033	0.376
Adjusted R^2	0.013	0.032	0.375
Residual Std. Error	127.584 (df = 6746)	126.390 (df = 6740)	101.534 (df = 6735)
F Statistic	46.090*** (df = 2; 6746)	28.492*** (df = 8; 6740)	312.479*** (df = 13; 6735)
Note:			*p<0.1; **p<0.05; ***p<0.01

TABLE B.10: Regression Model, Dep. Var.: Average Distance (Shannon Entropy) - B

	Dependent variable:		
	Average Ditance		
	(4)	(5)	(6)
n	0.759***	0.851***	0.787***
	(0.128)	(0.130)	(0.128)
Shannon_entropia	32.599***	29.097***	31.870***
	(2.550)	(2.726)	(2.562)
nord_ovest	−275.998***		
	(5.214)		
nord_est	−290.450***		
	(5.078)		
centro	−288.802***		
	(5.105)		
sud	−273.136***		
	(5.266)		
isole	−258.064***		
	(7.093)		
d_agr		1.756	
		(5.046)	
d_altro		21.580	
		(18.103)	
d_comm		3.645	
		(5.985)	
d_ind		17.935***	
		(4.264)	
d_serv		26.126***	
		(4.155)	
d_tur		9.042	
		(8.870)	
d_piemonte		−281.983***	−285.256***
		(6.887)	(6.845)
d_valle_d_aosta		−314.703***	−322.312***
		(19.687)	(19.725)
d_lombardia		−268.401***	−266.497***
		(5.674)	(5.653)
d_trentino		−296.268***	−296.711***
		(10.156)	(10.172)
d_veneto		−288.567***	−289.038***
		(5.900)	(5.867)
d_friuli		−296.914***	−300.399***
		(7.169)	(7.098)
d_liguria		−298.103***	−300.515***
		(10.175)	(10.173)
d_emilia		−285.484***	−284.309***
		(6.336)	(6.307)
d_toscana		−288.370***	−289.854***
		(6.541)	(6.485)
d_umbria		−286.717***	−288.348***
		(10.154)	(10.142)
d_marche		−297.776***	−296.068***
		(8.359)	(8.341)
d_lazio		−287.184***	−286.493***
		(5.729)	(5.678)
d_abruzzo		−296.520***	−297.578***
		(7.574)	(7.556)
d_molise		−289.059***	−290.335***
		(27.437)	(27.545)
d_campania		−272.363***	−272.646***
		(6.540)	(6.535)
d_puglia		−257.126***	−258.220***
		(6.824)	(6.846)
d_basilicata		−269.931***	−274.294***
		(15.180)	(15.200)
d_calabria		−262.896***	−262.416***
		(10.690)	(10.697)
d_sicilia		−246.573***	−248.235***
		(8.289)	(8.294)
d_sardegna		−270.339***	−273.803***
		(9.970)	(9.961)
Constant	307.182***	293.703***	307.465***
	(4.583)	(5.444)	(4.570)
Observations	6,749	6,749	6,749
R^2	0.369	0.381	0.375
Adjusted R^2	0.368	0.378	0.372
Residual Std. Error	102.086 (df = 6741)	101.270 (df = 6720)	101.740 (df = 6726)
F Statistic	562.805*** (df = 7; 6741)	147.628*** (df = 28; 6720)	183.071*** (df = 22; 6726)
Note:			*p<0.1; **p<0.05; ***p<0.01

TABLE B.11: Regression Results, Dep. Var.: Average Distance (Normalised Entropy) - A

	Dependent variable:		
	Average Distance		
	(7)	(8)	(9)
n	0.652^{***}	0.821^{***}	1.139^{***}
	(0.156)	(0.156)	(0.127)
H_norm	36.205^{***}	19.543^{***}	18.889^{***}
	(3.887)	(4.527)	(3.158)
d_agr		-44.185^{***}	
		(6.833)	
d_altro		-1.548	
		(22.565)	
d_comm		-26.027^{***}	
		(7.785)	
d_ind		-8.151	
		(5.557)	
d_serv		9.206^{*}	
		(5.445)	
d_tur		-15.451	
		(11.282)	
nord_ovest			-271.272^{***}
			(5.267)
nord_est			-286.909^{***}
			(5.147)
centro			-284.421^{***}
			(5.173)
sud			-268.863^{***}
			(5.324)
isole			-251.395^{***}
			(7.153)
Constant	49.854^{***}	65.717^{***}	310.894^{***}
	(2.669)	(6.082)	(4.881)
Observations	6,749	6,749	6,749
R^2	0.014	0.032	0.357
Adjusted R^2	0.014	0.031	0.356
Residual Std. Error	127.526 (df = 6746)	126.445 (df = 6740)	103.043 (df = 6741)
F Statistic	49.191^{***} (df = 2; 6746)	27.738^{***} (df = 8; 6740)	534.601^{***} (df = 7; 6741)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.12: Regression Model, Dep. Var.: Average distance (Normalised Entropy) - B

	Dependent variable:		
	Average Distance		
	(10)	(11)	(12)
n	0.759***	1.207***	1.161***
	(0.128)	(0.127)	(0.127)
Shannon_entropia	32.599***		
	(2.550)		
nord_ovest	−275.998***		
	(5.214)		
nord_est	−290.450***		
	(5.078)		
centro	−288.802***		
	(5.105)		
sud	−273.136***		
	(5.266)		
isole	−258.064***		
	(7.093)		
H_norm		15.325***	17.358***
		(3.674)	(3.176)
d_agr		−3.649	
		(5.584)	
d_altro		24.844	
		(18.327)	
d_comm		3.420	
		(6.319)	
d_ind		20.068***	
		(4.555)	
d_serv		27.775***	
		(4.428)	
d_tur		13.020	
		(9.141)	
d_piemonte		−277.324***	−281.570***
		(6.923)	(6.920)
d_valle_d_aosta		−308.385***	−317.574***
		(19.817)	(19.904)
d_lombardia		−263.732***	−261.241***
		(5.696)	(5.698)
d_trentino		−294.945***	−296.170***
		(10.234)	(10.288)
d_veneto		−285.285***	−285.895***
		(5.942)	(5.946)
d_friuli		−293.631***	−298.824***
		(7.215)	(7.189)
d_liguria		−294.880***	−298.586***
		(10.249)	(10.289)
d_emilia		−280.632***	−279.027***
		(6.362)	(6.360)
d_toscana		−284.180***	−286.321***
		(6.581)	(6.571)
d_umbria		−281.037***	−282.621***
		(10.213)	(10.238)
d_marche		−294.601***	−292.379***
		(8.419)	(8.433)
d_lazio		−282.698***	−282.212***
		(5.756)	(5.743)
d_abruzzo		−290.340***	−291.320***
		(7.603)	(7.619)
d_molise		−283.188***	−284.271***
		(27.627)	(27.797)
d_campania		−269.505***	−270.223***
		(6.593)	(6.623)
d_puglia		−253.496***	−254.474***
		(6.863)	(6.902)
d_basilicata		−260.916***	−266.212***
		(15.263)	(15.335)
d_calabria		−258.569***	−258.427***
		(10.770)	(10.824)
d_sicilia		−240.116***	−241.741***
		(8.326)	(8.365)
d_sardegna		−263.639***	−267.517***
		(10.019)	(10.048)
Constant	307.182***	296.916***	311.776***
	(4.583)	(6.278)	(4.870)
Observations	6,749	6,749	6,749
R^2	0.369	0.372	0.363
Adjusted R^2	0.368	0.369	0.361
Residual Std. Error	102.086 (df = 6741)	101.993 (df = 6720)	102.676 (df = 6726)
F Statistic	562.805*** (df = 7; 6741)	142.153*** (df = 28; 6720)	174.199*** (df = 22; 6726)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.13: Regression Model, Dep. Var.: log(*Average Distance*) (Shannon Entropy) - A

	Dependent variable:		
	log(*Average Distance*)		
	(1)	(2)	(3)
n	0.001	0.006^{***}	0.009^{***}
	(0.002)	(0.002)	(0.002)
Shannon_entropia	0.352^{***}	0.195^{***}	0.368^{***}
	(0.039)	(0.040)	(0.036)
d_agr		-0.750^{***}	-0.202^{***}
		(0.079)	(0.071)
d_altro		0.432	0.618^{*}
		(0.367)	(0.322)
d_comm		-0.390^{***}	0.003
		(0.097)	(0.086)
d_ind		-0.255^{***}	0.116^{**}
		(0.066)	(0.059)
d_serv		0.142^{**}	0.419^{***}
		(0.066)	(0.058)
d_tur		−0.231	0.114
		(0.141)	(0.125)
nord_ovest			-2.145^{***}
			(0.065)
nord_est			-2.241^{***}
			(0.063)
centro			-2.141^{***}
			(0.065)
sud			-2.141^{***}
			(0.066)
isole			-1.949^{***}
			(0.093)
Constant	3.418^{***}	3.662^{***}	5.136^{***}
	(0.032)	(0.064)	(0.068)
Observations	5,006	5,006	5,006
R^2	0.017	0.060	0.277
Adjusted R^2	0.017	0.059	0.275
Residual Std. Error	1.387 (df = 5003)	1.357 (df = 4997)	1.191 (df = 4992)
F Statistic	43.494^{***} (df = 2; 5003)	40.101^{***} (df = 8; 4997)	147.263^{***} (df = 13; 4992)
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

TABLE B.14: Regression Model, Dep. Var.: log(*Average Distance*) (Shannon Entropy) - B

	Dependent variable:		
	log(*Average Distance*)		
	(4)	(5)	(6)
n	0.006***	0.009***	0.007***
	(0.002)	(0.002)	(0.002)
Shannon_entropia	0.452***	0.361***	0.433***
	(0.034)	(0.036)	(0.034)
nord_ovest	−2.195***		
	(0.065)		
nord_est	−2.302***		
	(0.063)		
centro	−2.179***		
	(0.065)		
sud	−2.168***		
	(0.066)		
isole	−2.022***		
	(0.094)		
d_agr		−0.158**	
		(0.071)	
d_altro		0.557*	
		(0.320)	
d_comm		−0.007	
		(0.085)	
d_ind		0.127**	
		(0.059)	
d_serv		0.397***	
		(0.058)	
d_tur		0.143	
		(0.124)	
d_piemonte		−2.212***	−2.315***
		(0.091)	(0.091)
d_valle_d_aosta		−3.222***	−3.444***
		(0.264)	(0.266)
d_lombardia		−2.030***	−2.035***
		(0.070)	(0.070)
d_trentino		−2.459***	−2.488***
		(0.145)	(0.147)
d_veneto		−2.302***	−2.339***
		(0.075)	(0.075)
d_friuli		−2.338***	−2.460***
		(0.092)	(0.091)
d_liguria		−2.701***	−2.805***
		(0.144)	(0.145)
d_emilia		−2.081***	−2.106***
		(0.081)	(0.080)
d_toscana		−2.272***	−2.337***
		(0.084)	(0.084)
d_umbria		−2.080***	−2.164***
		(0.135)	(0.136)
d_marche		−2.416***	−2.437***
		(0.108)	(0.108)
d_lazio		−1.981***	−1.984***
		(0.076)	(0.076)
d_abruzzo		−2.413***	−2.464***
		(0.101)	(0.102)
d_molise		−2.194***	−2.263***
		(0.378)	(0.382)
d_campania		−2.220***	−2.236***
		(0.083)	(0.084)
d_puglia		−1.952***	−1.951***
		(0.087)	(0.088)
d_basilicata		−1.816***	−1.929***
		(0.207)	(0.209)
d_calabria		−2.011***	−2.027***
		(0.140)	(0.141)
d_sicilia		−1.866***	−1.905***
		(0.113)	(0.114)
d_sardegna		−2.070***	−2.163***
		(0.126)	(0.126)
Constant	5.288***	5.144***	5.297***
	(0.055)	(0.068)	(0.054)
Observations	5,006	5,006	5,006
R^2	0.254	0.291	0.273
Adjusted R^2	0.253	0.287	0.269
Residual Std. Error	1.209 (df = 4998)	1.181 (df = 4977)	1.196 (df = 4983)
F Statistic	243.186*** (df = 7; 4998)	73.023*** (df = 28; 4977)	84.930*** (df = 22; 4983)
Note:			*p<0.1; **p<0.05; ***p<0.01

TABLE B.15: Regression Model, Dep. Var.: log(*Average Distance*) Normalised Entropy - A

	Dependent variable:		
		log(*Average Distance*	
	(7)	(8)	(9)
n	0.006^{***}	0.009^{***}	0.011^{***}
	(0.002)	(0.002)	(0.002)
H_norm	0.538^{***}	0.314^{***}	0.385^{***}
	(0.051)	(0.058)	(0.046)
d_agr		-0.628^{***}	
		(0.087)	
d_altro		0.560	
		(0.368)	
d_comm		-0.273^{***}	
		(0.101)	
d_ind		-0.152^{**}	
		(0.070)	
d_serv		0.239^{***}	
		(0.069)	
d_tur		-0.087	
		(0.144)	
nord_ovest			-2.093^{***}
			(0.065)
nord_est			-2.213^{***}
			(0.064)
centro			-2.074^{***}
			(0.065)
sud			-2.070^{***}
			(0.067)
isole			-1.886^{***}
			(0.094)
Constant	3.340^{***}	3.519^{***}	5.264^{***}
	(0.036)	(0.078)	(0.060)
Observations	5,006	5,006	5,006
R^2	0.022	0.061	0.239
Adjusted R^2	0.022	0.060	0.238
Residual Std. Error	1.383 (df = 5003)	1.356 (df = 4997)	1.221 (df = 4998)
F Statistic	57.510^{***} (df = 2; 5003)	40.890^{***} (df = 8; 4997)	223.862^{***} (df = 7; 4998)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table B.16: Regression Model, Dep. Var.: log(*Average Distance*) (Normalised Entropy) - B

	Dependent variable:		
	log(*Average Distance*)		
	(10)	(11)	(12)
n	0.006***	0.012***	0.011***
	(0.002)	(0.002)	(0.002)
Shannon_entropia	0.452***		
	(0.034)		
nord_ovest	−2.195***		
	(0.065)		
nord_est	−2.302***		
	(0.063)		
centro	−2.179***		
	(0.065)		
sud	−2.168***		
	(0.066)		
isole	−2.022***		
	(0.094)		
H_norm		0.287***	0.350***
		(0.051)	(0.046)
d_agr		−0.165**	
		(0.078)	
d_altro		0.650**	
		(0.323)	
d_comm		0.041	
		(0.090)	
d_ind		0.192***	
		(0.062)	
d_serv		0.455***	
		(0.061)	
d_tur		0.255**	
		(0.127)	
d_piemonte		−2.125***	−2.217***
		(0.092)	(0.092)
d_valle_d_aosta		−3.129***	−3.358***
		(0.266)	(0.269)
d_lombardia		−1.959***	−1.938***
		(0.070)	(0.070)
d_trentino		−2.422***	−2.439***
		(0.146)	(0.148)
d_veneto		−2.237***	−2.253***
		(0.075)	(0.076)
d_friuli		−2.275***	−2.398***
		(0.092)	(0.092)
d_liguria		−2.626***	−2.727***
		(0.145)	(0.147)
d_emilia		−2.005***	−2.004***
		(0.081)	(0.081)
d_toscana		−2.196***	−2.242***
		(0.085)	(0.085)
d_umbria		−1.984***	−2.043***
		(0.136)	(0.137)
d_marche		−2.352***	−2.347***
		(0.108)	(0.109)
d_lazio		−1.901***	−1.885***
		(0.076)	(0.077)
d_abruzzo		−2.310***	−2.334***
		(0.101)	(0.102)
d_molise		−2.094***	−2.139***
		(0.380)	(0.385)
d_campania		−2.161***	−2.161***
		(0.084)	(0.085)
d_puglia		−1.894***	−1.880***
		(0.088)	(0.089)
d_basilicata		−1.666***	−1.764***
		(0.208)	(0.211)
d_calabria		−1.928***	−1.921***
		(0.141)	(0.143)
d_sicilia		−1.751***	−1.764***
		(0.114)	(0.115)
d_sardegna		−1.973***	−2.049***
		(0.126)	(0.128)
Constant	5.288***	5.096***	5.286***
	(0.055)	(0.080)	(0.060)
Observations	5,006	5,006	5,006
R^2	0.254	0.281	0.258
Adjusted R^2	0.253	0.277	0.255
Residual Std. Error	1.209 (df = 4998)	1.189 (df = 4977)	1.208 (df = 4983)
F Statistic	243.186*** (df = 7; 4998)	69.557*** (df = 28; 4977)	78.735*** (df = 22; 4983)
Note:			*p<0.1; **p<0.05; ***p<0.01

Bibliography

[Agarwal and Mittal, 2012] Agarwal, B. and Mittal, N. (2012). Hybrid approach for detection of anomaly network traffic using data mining techniques. *Procedia Technology*, 6:996–1003.

[Al-Rodhan and Stoudmann, 2006] Al-Rodhan, N. R. and Stoudmann, G. (2006). Definitions of globalization: A comprehensive overview and a proposed definition. *Program on the geopolitical implications of globalization and transnational security*, 6(1-21).

[Antonietti et al., 2019] Antonietti, R., Burlina, C., et al. (2019). From variety to economic complexity: Empirical evidence from italian regions. Technical report, Utrecht University, Department of Human Geography and Spatial Planning.

[Aureli and Del Baldo, 2014] Aureli, S. and Del Baldo, M. (2014). Contratto di rete e imprese artigiane: quali prospettive per l'internazionallizzazione? *Quaderni di ricerca sull'artigianato*, 2(1):21–48.

[Bank of Italy, 2020] Bank of Italy (2020). Investimenti diretti per paese controparte. https://www.bancaditalia.it/statistiche/tematiche/rapporti-estero/investimenti-diretti/index.html.

[Bartoli and Rizzi, 2017] Bartoli, M. and Rizzi, E. (2017). Le regioni a favore delle reti d'impresa - studio sui finanziamenti per le aggregazioni. https://www.retimpresa.it/le-regioni-a-favore-delle-reti-di-impresa/.

[Bernard et al., 2007] Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2007). Firms in international trade. *Journal of Economic perspectives*, 21(3):105–130.

[Borra and Di Ciaccio, 2014] Borra, S. and Di Ciaccio, A. (2014). *Statistica: metodologie per le scienze economiche e sociali.* McGraw-Hill.

[Brandolini and Bugamelli, 2009] Brandolini, A. and Bugamelli, M. (2009). Rapporto sulle tendenze nel sistema produttivo italiano. *Questioni di economia e finanza*, 45.

[Bronzini, 2007] Bronzini, R. (2007). FDI Inflows, agglomeration and host country firms' size: Evidence from Italy. *Regional studies*, 41(7):963–978.

[Buczkowska et al., 2019] Buczkowska, S., Coulombel, N., and de Lapparent, M. (2019). A comparison of euclidean distance, travel times, and network distances in location choice mixture models. *Networks and spatial economics*, 19(4):1215–1248.

[Burlina, 2020] Burlina, C. (2020). Networking policy and firm performance. *Growth and Change*, 51(1):161–179.

[Cabigiosu, 2021a] Cabigiosu, A. (2021a). Caratteristiche, nascita ed evoluzione dei contratti di rete. *Osservatorio Nazionale sulle reti d'impresa 2021*, page 35.

[Cabigiosu, 2021b] Cabigiosu, A. (2021b). Osservatorio nazionale sulle reti d'impresa 2021. *Osservatorio Nazionale sulle reti d'impresa 2021*.

[Camere di Commercio d'Italia, 2022] Camere di Commercio d'Italia (2022). Le statistiche. https://contrattidirete.registroimprese.it/reti/index.action. Accessed: 2022-07-03.

[Cassa Depositi e Prestiti, 2018] Cassa Depositi e Prestiti (2018). Il sistema produttivo italiano. tra modernizzazione e industria 4.0.

[Cerutti and De Falco, 2020] Cerutti, S. and De Falco, S. (2020). Il ruolo della prossimità nella geografia dell'impresa. un'analisi comparata plurisettoriale tra le città di napoli e novara. *Bollettino dell'Associazione Italiana di Cartografia*.

[Comin, 2010] Comin, D. (2010). Total factor productivity. In *Economic growth*, pages 260–263. Springer.

[Costa et al., 2017] Costa, S., Lucchetti, F., and Livia, R. (2017). Gli effetti sul contratto di rete sulla performance delle imprese.

[Di Dio and Correani, 2020] Di Dio, F. and Correani, L. (2020). Quality-improving and cost-reducing strategic alliances. *Economia Politica*, 37(2):493–524.

[Di Dio et al., 2022] Di Dio, F., Correani, L., and Morganti, P. (2022). Does the nature of goods affect bilateral exchange of technology and location choice in stable networks? *Journal of Industry, Competition and Trade*, pages 1–15.

[Di Giacinto et al., 2022] Di Giacinto, V., Secchi, A., and Tosoni, A. (2022). Nuova vita per i distretti. https://www.lavoce.info/archives/97071/nuova-vita-per-i-distretti/.

[Economides, 1996] Economides, N. (1996). The economics of networks. *International journal of industrial organization*, 14(6):673–699.

[Eurostat, 2018] Eurostat (2018). Glossary: Foreign Direct Investment (FDI). https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Foreign_direct_investment_(FDI)#:~:text=Foreign%20direct%20investment%2C%20abbreviated%20as,enterprise%20resident%20in%20another%20economy.

[Fabrizi et al., 2022] Fabrizi, A., Guarini, G., and Meliciani, V. (2022). Environmental networks and employment creation: Evidence from italy. *Journal of Cleaner Production*, 359:132056.

[Garofalo, 2007] Garofalo, G. (2007). *Capitalismo distrettuale, localismi d'impresa, globalizzazione*. Firenze University Press.

[Garofalo and Pugliesi, 2014] Garofalo, G. and Pugliesi, S. (2014). *"Fare rete": la propensione delle imprese alla cooperazione in un contesto locale*. Narcissus Accademy.

[Hulten, 2001] Hulten, C. R. (2001). Total factor productivity: a short biography. In *New developments in productivity analysis*, pages 1–54. University of Chicago Press.

[Istat, 2022a] Istat (2022a). Ateco: Struttura, Contenuti, Criteri. https://www.istat.it/it/archivio/266175. Accessed: 2022-11-20.

[Istat, 2022b] Istat (2022b). Struttura ateco 2007 aggiornamento 2022. https://www.istat.it/it/archivio/266993,.

[Krugman and Obstfeld, 2009] Krugman, P. R. and Obstfeld, M. (2009). *International economics: Theory and policy*. Pearson Education.

[Marshall, 2009] Marshall, A. (2009). *Principles of economics: unabridged eighth edition*. Cosimo, Inc.

[Milner, 1999] Milner, H. V. (1999). The political economy of international trade. *Annual review of political science*, 2(1):91–114.

[Mori and Nishikimi, 2002] Mori, T. and Nishikimi, K. (2002). Economies of transport density and industrial agglomeration. *Regional Science and Urban Economics*, 32(2):167–200.

[Napolitano et al., 2018] Napolitano, O., Pietroluongo, M., Kounetas, K., et al. (2018). Stochastic convergence or divergence of total factor productivity and gdp of italian regions. re-examing the evidence. *Economics Bulletin*, 38(4):1857–1863.

[Pastore et al., 2020] Pastore, P., Ricciardi, A., and Tommaso, S. (2020). Contractual networks: an organizational model to reduce the competitive disadvantage of small and medium enterprises (smes) in europe's less developed regions. a survey in southern italy. *International Entrepreneurship and Management Journal*, 16(4):1503–1535.

[Pebesma, 2018] Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1):439–446.

[Pitingaro and Corsini, 2021] Pitingaro, S. and Corsini, S. (2021). I contratti di rete: un'analisi descrittiva attraverso i dati del registro imprese. *Osservatorio Nazionale sulle reti d'impresa 2021*, page 15.

[Proto, 2021] Proto, A. (2021). L'evoluzione delle relazioni banche-reti. *Osservatorio Nazionale sulle reti d'impresa 2021*, page 53.

[RetImpresa, 2019a] RetImpresa (2019a). Report sulle reti di imprese in italia - 1° semetre 2019. https://www.retimpresa.it/report-sulle-reti-di-imprese-in-italia-i-semestre-2019-2/.

[RetImpresa, 2019b] RetImpresa (2019b). Report sulle reti di imprese in italia 2018. https://www.retimpresa.it/report-annuale-2018/.

[RetImpresa, 2020] RetImpresa (2020). Report sulle reti di imprese in italia - 1° semetre 2020. https://www.retimpresa.it/report-reti-dimpresa-1-semestre-2020/.

[RetImpresa, 2021] RetImpresa (2021). Il 2020 in rete. https://www.retimpresa.it/report/.

[Rizzo, 2020] Rizzo, M. C. (2020). Le reti di impresa: modelli innovativi di business post covid-19. https://www.altalex.com/documents/news/2020/04/24/reti-di-impresa-modelli-innovativi-di-business-post-covid-19. Accessed: 2022-11-29.

[Romano et al., 2016] Romano, L., Paolazzi, L., De Vita, L., Costa, S., and Lucchetti, F. (2016). Reti d'impresa. l'identikit di chi si aggrega: competitivo e orientato ai mercati esteri. *Centro Studi Confindustria RetImpresa*.

[Sciarelli and Tani, 2014] Sciarelli, M. and Tani, M. (2014). La social network analysis per lo studio dell'innovazione nelle reti di imprese. *Sinergie quaderni di ricerca*, (17):87–104.

[Shannon, 1951] Shannon, C. E. (1951). Prediction and entropy of printed english. *Bell system technical journal*, 30(1):50–64.

[Shannon, 2001] Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review*, 5(1):3–55.

[Shimizu, 2002] Shimizu, D. (2002). Product differentiation in spatial cournot markets. *Economics Letters*, 76(3):317–322.

[Signorell et mult. al., 2022] Signorell et mult. al., A. (2022). DescTools: Tools for descriptive statistics. https://cran.r-project.org/package=DescTools,. R package version 0.99.47.

[University of Groningen and University of California, Davis, 2022a] University of Groningen and University of California, Davis (2022a). Real Gross Domestic Product for Italy [CLVMNACNSAB1GQIT]. https://fred.stlouisfed.org/series/CLVMNACNSAB1GQIT. Accessed: 2022-12-15.

[University of Groningen and University of California, Davis, 2022b] University of Groningen and University of California, Davis (2022b). Total Factor Productivity at Constant National Prices for Italy [RTFPNAITA632NRUG]. https://fred.stlouisfed.org/series/RTFPNAITA632NRUG. Accessed: 2022-11-28.

[Vătămănescu et al., 2020] Vătămănescu, E.-M., Alexandru, V.-A., Mitan, A., and Dabija, D.-C. (2020). From the deliberate managerial strategy towards international business performance: A psychic distance vs. global mindset approach. *Systems Research and Behavioral Science*, 37(2):374–387.

[Vendetti, 2022] Vendetti, A. (2022). III trimestre 2022 - esportazioni delle rerioni italiane. https://www.istat.it/it/archivio/278620#:~:text=Nel%20terzo%20trimestre%202022%2C%20si,Isole%20(%2D0%2C6%25).. Accessed: 2022-04-01.

[Von Mises, 2014] Von Mises, R. (2014). *Mathematical theory of probability and statistics*. Academic press.

[Wanzenböck, 2018] Wanzenböck, I. (2018). A concept for measuring network proximity of regions in r&d networks. *Social Networks*, 54:314–325.

[World Bank Group, 2022] World Bank Group (2022). GDP growth (annual %) - Italy. https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?end=2021&locations=IT&start=1961. Accessed: 2022-11-28.

[Zazzaro, 2010] Zazzaro, A. (2010). Reti d'imprese e territorio: tra vincoli e nuove opportunità dopo la crisi. *Il Mulino Editore, Bologna*.

[Zecca et al., 2014] Zecca, F., Al Am, A., and Capocchi, E. (2014). Dai distretti alle reti d'impresa: soluzioni chiave per lo sviluppo territoriale. *Italian Review of Agricultural Economics*, 69(2-3):227–243.