Ca' Foscari
University
of Venice

**Department
of Economics**

Working Paper

Carlo Gianelle

**Exploring the complex structure
of labour mobility networks.
Evidence from Veneto microdata**

ISSN: 1827/3580
No. 13/WP/2011

Exploring the complex structure of labour mobility networks. Evidence from Veneto microdata

Carlo Gianelle
University of Venice

First Draft: June 2011

Abstract

This essay investigates the network structure of inter-firm worker mobility in Veneto, an industrial region of Northern Italy, using comprehensive employer-employee matched data. The empirical network reveals a small world pattern that hinges critically upon a few hub firms. Main hubs are found to be: (1) long-established manufacturing companies; (2) wholesale companies; and (3) companies supplying workforce to third parties. The methodology of investigation provides a toolkit for monitoring labour market evolution, and should enable industry policies supporting labour reallocation mechanisms.

Keywords

Regional labour markets, worker reallocation, complex networks, small world, hub dependence

JEL Codes

D85, J63, R12

Address for correspondence:

Carlo Gianelle

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
Phone: (+39) 041 234173
Fax: (+39) 041 2349176
e-mail: carlo.gianelle@unive.it

This Working Paper is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.

1. Introduction

This paper investigates the network structure of inter-firm worker mobility in a regional labour market in Northern Italy during the 1990s. Worker reallocations are selected and assembled together so as to form a knit texture, formally defined as a binary directed graph where vertices indicate firms and links represent transfers of workers between firms; the characteristics of such structure are examined using the original lens of complex network theory.

The study represents a tentative bridging between the analyses of labour mobility based on linked employer-employee data and the analyses of complex networked systems applied to human behaviours. The main goal is to improve the general understanding of the worker reallocation phenomenon in a well-defined regional labour market, offering policy makers a new tested toolkit for monitoring worker flows that could support precisely targeted interventions directed at improving or preserving the functioning of reallocation mechanisms.

For the first time, a large worker mobility network, consisting of hundreds of thousands of vertices and links and covering the universe of private sector employment in an industrial region of Europe over a full decade, is thoroughly analyzed. In such a context, network analysis allows to explore several aspects of labour mobility that otherwise would remain hidden in the mass of individual work histories contained and intertwined in a linked employer-employee dataset, and which could not be easily mined by standard econometrics. In particular, two are the issues it is possible to appreciate: whether the regional economy under scrutiny proves to be globally integrated into a single labour market and to what extent this labour market is easily practicable by workers; and

whether or not the flowing of workers from job to job depends critically on a few pivotal employers.

The investigation is based upon mapping individual reallocations in a region of Italy, Veneto. The key statistical source is the VWH database, a longitudinal dataset that covers the universe of worker histories in Veneto in the private sector for more than twenty years, up to 2001.¹ In respect to the specific research purposes of this essay, the VWH database presents the unique feature of allowing to recover virtually all individual transitions between firms; moreover, the administrative nature of the data ensures that the obtained network is indeed a reliable representation of the true fabric of labour mobility.

Veneto is a highly industrialized Nuts-2 region that ranks among the most developed and densely populated regions of Europe. Most notably, the economy is by far dominated by small and medium size ventures; economic activities are widespread over the entire territory, at the same time showing clear spatial agglomeration pivoted around a plurality of small and medium urban centres and industrial districts (Brusco, 1986; Becattini et al., 1990), but overall there emerges no single organization nor economic concentration that can play as a gravity centre of the system, contrary for instance to regions hosting a big metropolitan area or deeply dependent on very large enterprises (Tattara and Anastasia, 2003). Throughout the period under scrutiny, the Veneto labour market has been characterized by a positive rate of job creation in both manufacturing and services, and by almost frictional

¹ VWH is the acronym for Veneto Worker Histories, a longitudinal linked employer-employee dataset developed at the Department of Economics of the University of Venice Ca' Foscari on the basis of administrative records of the Italian Social Security Institute.

unemployment, accompanied by high labour mobility, with worker turnover comparable to Anglo-Saxon countries.

Well understood network models are applied in order to characterize the structure of inter-firm labour mobility. The first focus is on assessing the extent to which the network is pervious to reallocation flows, allowing workers the potentiality to navigate easily across the labour market; to this aim the actual arrangement of inter-firm links is studied and it is compared with the ‘small world’ model (Milgram, 1967; Watts, 1999; Newman, 2000; Goyal et al., 2006), which is taken as a benchmark for efficient mobility (Latora and Marchiori, 2001). The attention is then turned to the role of employers, examining the statistical distribution of firm links, in order to appreciate how connectivity is distributed within the system; to this aim the empirical connectivity distribution is compared with Pareto-like models (Barabasi and Albert, 1999; Newman, 2005; Gabaix, 2009). Finally, the most connected firms are closely examined, appreciating their importance in integrating different areas of the regional mobility network by means of evaluating the consequences of simulating the failure of such firms, as opposed to the failure of firms chosen at random (Albert et al., 2000).

The paper is organized as follows. Section two briefly discusses the nexus between labour reallocation and networks; Section three formally introduces the network concepts and the empirical strategy; Section four illustrates the data and the main traits of the Veneto economy; Section five contains the small-world analysis; Section six focuses on the distribution of links and on the role of most connected firms; Section seven highlights the characteristics of important employers; finally, Section eight concludes.

2. Worker reallocation and network approach

The increasing availability of longitudinal linked employer-employee data (Abowd and Kramarz, 1999; Hamermesh, 2008) and the recent development of interpretative models of complex networked systems (Watts, 2003; Newman et al., 2006; Jackson, 2008) offer a unique chance of analyzing in a systematic way the entangled set of tracks naturally arising from the flowing of workers between employers in a given territory.

A distinguishing feature of longitudinal linked employer-employee data is that individuals and their employers are identified and followed over time. Labour economists have taken full advantage of these data to revisit several classic questions, related to employment and wage dynamics, labour turnover, reallocation and productivity growth, just to name a few. More importantly, the endeavour in the construction and analysis of these data has shown its potential in opening a host of new research directions, often at the crossroads between different branches of economics and sociology. In particular, the powerful support of linked employer-employee data allows to scrutinize anew the micro-level behaviour of worker flows originated by individual reallocations; the progress in the science of networks offers the analytical tools to effectively mine information from the data.

The interest in seeking for new aspects of labour mobility finds its root in the key role labour reallocation plays in most industrial economies, with critical implications both on individual working lives, and on the productivity of the economic system. The reallocation of production factors between concurrent uses is widely understood to represent an

important ingredient of growth in market economies.² In real markets, firms, procedures, and products are continuously replaced by new ones, while capital is diverted between alternative investments, and workers transfer across firms, much in the Schumpeterian spirit of creative destruction.

Most dynamic economies exhibit significant rates of resource reallocation across production units and economic sectors; in particular, the reallocation of workers between jobs in the context of regional economies emerges as a distinctive feature of industrial economies. Not only labour mobility is essential for a system in order to adapt and respond quickly and effectively to competitive and fast-changing international scenarios, but it may significantly contribute to the spread of knowledge, ideas, and know-how within the economic system, possibly resulting in higher productivity and growth.³

Regional economies which are administratively bounded represent an appropriate setting for empirically investigating labour mobility, which tends to be mostly localized and it is influenced by the institutional environment. In such contexts, the traditional flow approach to labour markets helps us studying the behaviour of aggregated employment flows according to cyclical fluctuations and/or sectoral and employer size classifications, but still it does not reveal the details of the inter-firm arrangement of job to job mobility flows.⁴

In this respect, the seminal work of Granovetter (1973) first addressed the problem of job changes at the individual level, using network-based concepts and showing the

² The positive relationship between reallocation and productivity has been addressed among others by Baily et al. (1992); Bartelsman and Dhrymes (1998); Foster et al. (2001); and more recently by Acemoglu (2009); empirical studies concerning transition Countries are those by Brown and Earle (2008).

³ The empirical regularities of these processes have been extensively investigated and documented, see for instance Davis and Haltiwanger (1999).

⁴ For a review of the flow approach to labour markets, see for instance Schettkat (1996), and Davis et al. (2006); a recent study of worker mobility in Veneto, using the flow approach, is that of Tattara and Valentini (2010).

importance of social ties or interpersonal connections in disclosing new job opportunities to workers. But analyses of the global architecture of firm to firm flows are lacking; in particular, the complex web of connections between firms each worker contributes to weave as the result of changing employer, has never been thoroughly investigated at large scale.⁵

At firm-level resolution, worker reallocation looks like a gradual process of percolation through a porous material, and the microscopic architecture of this material is what ultimately determines the direction and timing of the flows which are usually observed at some higher level of aggregation. The present essay adopts a network representation of inter-firm worker mobility – where firms denote the vertices and worker flows sedimented in time denote the links – that can usefully support original investigations on the functioning of the labour market.

3. Network concepts and empirical strategy

Two are the main issues the present study seeks to quantitatively address by means of a network approach: (1) understanding how much the reallocation market behaves like an integrated system that is pervious to labour flows; (2) identifying which are, if any, the pivotal employers acting as crossroads or brokers for reallocation flows, and what is their role in the global architecture of the network. For the network

⁵ Network methodologies have been recently spreading in many different fields of economics. Prominent examples can be found for international trade (Fagiolo et al., 2010), knowledge flows and patent citations (Sorenson et al., 2006; Fleming et al., 2007), and collaborations in research and innovation (Frenken, 2000; Fleming and Frenken, 2007; Verspagen and Duysters, 2007; Ter Wal, 2010; Kauffeld-Monz and Fritsch, 2010), just to name a few. To the best of our knowledge, the only work dealing with large labour mobility networks is that of Currarini and Feri (2006).

nomenclature to be used henceforth, see the Appendix, which provides precise mathematical definitions.

The first line of enquiry is developed in order to primarily appreciate whether the firms in the reallocation market prove to be integrated into a unique web of flows, that is a network with links arranged in such a way that vertices are interconnected to each other by means of uninterrupted sequences of links, called paths in network jargon, or the global network is constituted of many disjointed networks, or subnetworks, revealing a segregated market. The analysis then proceeds exploring whether the actual arrangement of links is such that the network distance between mutually reachable vertices, that is to say the number of contiguous links that one counts in order to get from one vertex to another, is appreciably small compared to the network size. This amounts to evaluate how much the most integrated portion of the network is effectively practicable, or traversable by worker flows in just a few steps, that is what is referred to as permeability or perviousness.

The assessment of integration and permeability of the market to labour flows may translate into important indications in support of policy decisions: for instance, an integrated and permeable market can in principle better absorb cyclical shocks manifesting locally, compared to a segregated market where shocks are more likely to produce long-term unemployment, possibly calling for interventions favouring re-employment.

The second line of inquiry aims at revealing the existence and understanding the role of particular employers that, for their position in the network and their prominence in terms of flows of reallocating workers both on entry and exit, act as key crossroads in the market, with crucial importance for both integration and permeability of the system. Such kinds of

players are the hubs in network jargon; many forms of integrated and permeable networks may exist (Amaral et al., 2000), the presence/absence of a hub backbone being a distinctive characteristic.

This analysis essentially exploits the network representation as a mapping device in order to pinpoint employers which are especially prominent in the reallocation market. Such monitoring may be relevant for policy, because hub employers of course have some power in terms of controlling and directing labour flows, potentially being of strategic importance for workers careers, and for the dissemination of competencies, practices, and know-how in the economy. Hence, hubs appears to be suitable target candidates for interventions aimed for instance at granting operation continuity, should they be deemed essential to a good functioning of the reallocation mechanisms.

The literature on complex network has been explored in search for appropriate ways of appreciating the network properties just discussed. Noticeably, network theory directly relates the notion of integration and permeability to a specific model of network known as the small world (Watts and Strogatz, 1998), whose crucial functional significance is precisely that of making vertices easily reachable one from another, employing a parsimonious assembly of links; importantly, the small-world model has been demonstrated to be highly efficient, in the sense that it grants reliable communication and flow exchange between vertices (Latora and Marchiori, 2001).⁶

The original approach adopted in this work hence uses the small-world properties of the inter-firm network as a proxy indicator of labour market integration and permeability to

⁶ For comprehensive reviews of the properties of small-world networks see for instance Newman (2000).

reallocation flows: the closer the actual network to a small world, the more integrated and permeable the system is to reallocation flows. According to a consolidated literature, a network is said to be a small world when the following four conditions simultaneously apply:

1. a large number of vertices is reachable from many other vertices in the network, i.e. there exist uninterrupted sequences of links connecting a relevant fraction of possible pairs of vertices;
2. vertices are reachable with little effort, i.e. interconnected vertices are on average only a few steps/links away from each other (where to go along a link is assumed to be costly);
3. the system is overall parsimonious, i.e. the actual number of links is much smaller than the maximum possible number, obtained when a link is placed between every possible pair of vertices (where to provide an additional link is assumed to be costly);
4. vertices form cohesive groups, i.e. vertex neighbours, or vertices directly connected with the given one, also tend to be interconnected among each others, meaning that there exist alternative ways of reaching a former neighbour vertex should the direct connection be lost.

The combination of such principles confers on the notions of integration and permeability an immediate economic meaning: the system is integrated and permeable if, whatever vertex is considered, both the number of vertices potentially reachable is high (condition 1), the cost of actually reaching such vertices is, on average, low (conditions 2 and 3), and the system is robust to local disconnection (condition 4). Often in

real-world situations people confront a positive relationship between the number and variety of alternative opportunities available for moves or choices, and the average cost of actually seizing such opportunities or reaching targets. The small world reveals itself to be an extremely effective configuration in order to ease such a problematic relationship, because it guarantees maximum reachability, while minimizing the costs both for the system designer – i.e. the cost of providing links – and for the system user – i.e. the cost of moving along sequences of links in order to reach a given target.

TABLE 1 AROUND HERE

The four defining properties of the small-world model can be re-framed in exact graph-theoretic terms with reference to a directed network, resulting in the four propositions listed in the second column of Table 1.

A small-world network comes into view as an integrated system essentially dominated by local clustering, with a relatively few long-range links that act as shortcuts, connecting different bunches of vertices which otherwise would be much farther away from each other. The giant strongly connected component guarantees the existence of connecting paths between all possible vertices comprised within its boundaries, while short distances indicate that vertices can be reached with little effort; besides, clustering – meaning redundancy of links at local level – promotes robustness to link disconnection and, through multiple independent pathways, reliable connectivity as well (White and Houseman, 2002).

In order to quantitatively appraise the small-world properties, evaluation criteria are derived from the leading

network literature and listed in the third column of Table 1. The combination of such criteria identifies a range of parameters values for which a network shows small-world characteristics. The first goal of the analysis thus becomes establishing whether the empirical labour mobility network is close to a small world or not.

The question to be addressed next is whether firms in the population under scrutiny tend to have similar positions in the network, or they exhibit some degree of heterogeneity. This amounts to check for the existence and to appreciate the role of pivotal employers in the reallocation market, that is employers which are very central in the network, being able of acting as major crossroads for reallocations. The most straightforward way of evaluating how much a firm is central in the reallocation network is to look at the number of links, the degree in network terminology, that the firm has both on entry and exit; this simple statistic is usually referred to as degree centrality (Freeman, 1979).

In particular, a way of grasping at a glance the centre-periphery structure of a network is to examine the statistical distribution of firm links. The presence of pivotal connectors is signalled by a right-skewed distribution, with a long or fat tail that contains vertices having many more connections than average and appearing much more frequently than in a standard Gaussian distribution. In the present study, the empirical strategy devised in order to verify the presence of a hub backbone consists of two steps: first analyzing some moments of the degree distribution, and then moving to the examination of the functional form of the statistical distribution.

The widely acknowledged literature on complex networks shows that the degree distribution of a network with a hub backbone is very often well approximated in the tail by a

Paretian distribution, more specifically, by a negative power law (Barabasi and Albert, 1999, 2002; Newman, 2003). This essay hence takes the power law distribution as a benchmark for evaluating the presence of a hub backbone; the empirical degree distribution is first fitted to a power law, then the goodness of the fit is rigorously evaluated using the tests proposed by Clauset et al. (2009).

A further phase of analysis aims at assessing the relevance of hubs in the global pattern of linking across firms. This test is performed by means of the technique proposed by Albert et al. (2000), consisting in computing the extent of network integration after simulating a series of attacks to hubs, and then comparing the results with those obtained when similar attacks are targeted to vertices chosen at random.

Finally, the main hub firms are identified and their activity, relevant characteristics, and historical background are discussed; a classification of important firms is then proposed and commented.

4. The Veneto economy and the VWH dataset

Inter-firm worker mobility is studied within an industrialized region of Northern Italy, Veneto, representing a well-defined labour market for which comprehensive and detailed information about worker careers are available over a long period of time. The actual analysis covers the period from 1991 to 2000. Basic statistics for the Veneto economy are listed in Table 2.

TABLE 2 AROUND HERE

Veneto is a highly industrialized NUTS2 region, ranking 58th among European regions in terms of per capita GDP in the year 2000, 31% above the EU-27 average (Eurostat, 2010); in the same year the total population amounts to nearly 4.5 millions, with a density of 256 inhabitants per square kilometer. Throughout the 1990s, the Veneto labour market has been characterized by nearly full employment, with an unemployment rate of 3% in the year 2000, and by a positive rate of job creation in almost all sectors of the economy.

As for industrial specialization, Veneto can be defined as a ‘manucentric’ region; the industrial system is characterized by a large population of small and medium firms, frequently organized in districts, whose historical specializations are garments, textiles, leather and shoes, goldsmiths, mechanical products, furniture, and plastics. The average firm size in 2001 was just 3 employees in the whole economy, 9 employees in manufacturing.

The basis for the network construction is the VWH panel, a matched employer-employee dataset derived from administrative records of the Italian Social Security Institute (Inps), referring to the entire population of private sector workers and employers in Veneto, during the period 1975-2001.⁷ The VWH dataset covers each single worker employed in the private sector and each single establishment with at least one employee. The information available allow to build a monthly history of the working life of each employee who has been hired for at least one day by an establishment based in Veneto, during the period of observation, regardless of the worker’s place of residence.

Firms can be distinguished through their univocal Tax Identification Number, while workers are marked by an

⁷ The VWH panel and the statistical procedures implemented in order to build it are documented in Occari et al. (2001).

anonymous individual code. In the source archives, the firm code changes each time the legal ownership of the firm changes, even if there is no stop in the underlying operations, nor appreciable change in the nature of the activity. Whenever such situation has been recognized, the business is said to be continuous, and the old employer (incorporated) is assigned the code of the new employer (incorporating).⁸

The labour mobility network is obtained by counting all the individual reallocations occurred within a given time window, where a reallocation is defined by an pair of events: a separation, i.e. the conclusion of a job relationship with a given employer, and a subsequent engagement of the same individual in a job relationship with another employer. Reallocations are then mapped onto a directed graph, where links represent flows of workers moving between two firms; the links represent the sedimentation of several individual reallocations between the same employers occurred over the years.

In the present essay, the empirical focus is restricted to reallocations involving workers of both sexes, aged between 15 and 65 years, whose separation occurred within the period 01/01/1991-12/31/2000, and where at most 12 months elapsed between the separation and the following engagement. No restrictions are imposed on the duration of job spells, nor on the type of occupation. Reallocations in the same firm are excluded; businesses not involved in labour reallocations do not enter the network, hence there are not isolated vertices in the resulting graph.

⁸ Formally, the legal employer A is considered incorporated into the legal employer B anytime more than 50% of its employees are taken over by B.

5. Small-world analysis

From the assembled data a labour mobility network of 379,391 vertices and 1,899,898 directed links is obtained; this is the main object of study. The basic network statistics are listed in Table 3.

Following the rules of Table 1, the first task in order to verify the consistency of the empirical pattern with the small-world one is to evaluate the level of network integration by looking at the network components. It can be noticed from the data in Table 3 that the network exhibits a giant weakly connected component comprising 98.6% of vertices; the WCC in turn exhibits an inner core given by a giant strongly connected component that covers 46.2% of vertices. At the same time, the size of the second largest weakly connected component and of the second largest strongly connected component are just 6 and 4 respectively, with a negligible role within the overall architecture. Noticeably, both the SCC and the WCC cover a much higher portion of vertices than required for the network to be a small world, hence the first condition of Table 1 is fulfilled.

TABLE 3 AROUND HERE

The investigation can be pushed further in order to uncover more subtle aspects of the relationship between the SCC and the WCC. Hence, a bow-tie type of analysis is performed (Broder et al., 2000), identifying two other groups of vertices belonging to the WCC and not to the SCC: an IN component, consisting of vertices that can only reach the SCC, and an OUT component, consisting of vertices that can only be reached from the SCC. In the actual network, the IN and OUT components together cover most of the WCC vertices which are

not directly part of the SCC; besides, 76% of the firms belonging to the IN and OUT components exhibits total degree equal to just one. Therefore, the topology of the network may be sketched as a huge flower eye of strongly interconnected firms, surrounded by a corolla of vertices, each of which have a single link, either pointing to, or coming from the eye.

Considering these results altogether, the first finding of the analysis can be therefore stated as follows: the labour mobility network shows a highly integrated configuration, with a centre of gravity in the SCC, which is compatible with the small-world model.

The actual link density of network scores $1.32e-05$, that is five orders of magnitude lower than the maximum possible density, that is the density of a complete network, where a link is present between each possible pair of vertices. The empirical value is very low, hence, the second finding of the analysis is that: the labour mobility network exhibits a sparse topology, as required by the small-world model.

A step further in the small-world analysis consists of computing the average network distance between vertices belonging to the SCC, deriving the average path length, and comparing it with the benchmark value $\ln(n_{scc})$. Table 3 reports a value of 4.41 for the APL, meaning the whole SCC – hence almost half of the global network, more than 175 000 vertices – can be traversed going along less than five links on average, which is a far small number compared to the SCC size, even smaller than the benchmark value $\ln(n_{scc})$. Such a result leads to conclude that: the labour mobility network shows the short distance typical of a small world.

Finally, in order to complete the small-world analysis the average clustering coefficient ACC is computed, so as to verify whether the labour mobility network exhibits an assembly of

cohesive subgroup of vertices into a broader architecture. The ACC reported in Table 3 is equal to $3.48e-02$; in order to evaluate this number, the actual network is compared with two randomized variants taken as null models (Watts and Strogatz, 1998; Newman et al. 2001): when there is no agreement between the actual network and the null models, a pattern in the data exists that is worth exploring; namely, if the actual network exhibits significantly higher clustering than the comparison networks, it can be rightly claim that vertices come into cohesive subgroups.

As first term of comparison, a network of the same number of vertices and links as the actual network, but with links placed completely at random, is built. Such network exhibits an ACC equal to $5.49e-05$, that is three orders of magnitude smaller than the one observed in the real data. Then, a second random network is built, of the same size as the actual network, and the links placed at random under the constraint that the resulting in-degree distribution must be equal to the actual one; such network yields an ACC equal to $3.53e-04$, that is two orders of magnitude smaller the actual value. Hence: the real network reveals itself to be appreciably clustered, as required by the small-world model.

The observed clusterization is likely to result from the combination of two factors: first, worker mobility tend to remain mostly local in spatial terms, because of the cost of long-distance commuting; second, economic activities tend to come in spatially localized groups, especially in Veneto, where Marshallian districts represent a major phenomenon. Districts are clusters of firms closely knit by means of a range of different types of relationships, among which the mobility of labour plays an important role (Tattara, 2001). Another distinguishing trait of Veneto is the presence of a plurality of

middle-size urban centres, with weak hierarchical relationships, which may deepen the tendency of labour mobility to exhibit clustering, since each conurbation is expected to be a densely interconnected labour market.

The general conclusion of the analysis performed in this Section is that: the labour mobility network in Veneto is a small world. The labour reallocation market consequently appears to be globally integrated and pervious to worker flows.

6. Hub analysis

The role of the firm populating the labour mobility network is explored by examining the statistical distribution of links. Table 3 displays an average total-degree – that is the average number of links per firm – of 10.02; the standard deviation associated to this indicator is 56.35, pointing to high variability in the link distribution. The total-degree sequence of the network ranges from 1 to 11643, with a median value of just 2; 82.5% of vertices has $k < k(G)$, 41.2% has $k=1$, and 1.3% has $k > 100$. Such a pattern points to a very unequal distribution, markedly right-skewed. Similar considerations hold true for both the in-degree and the out-degree distributions.

In Figure 1 the complementary cumulative distributions, CCDDs, for total-degree (a), in-degree (b), and out-degree (c) are plotted on a double logarithmic scale.⁹ It can be noticed at a glance that each plot follows a clear-cut negative relation that appears to be almost linear above some threshold, located at $k \approx 100$ for the total-degree and the in-degree, and at $k \approx 50$ for the out-degree. This pattern is characteristic of probability distributions belonging to the Paretian family, namely of power

⁹ The complementary cumulative distribution guarantees a better visual display than the basic frequency distribution, because it reduces possible fluctuations in the extreme right tail, due to the low number of observations in this region (Newman, 2005).

laws, which have attracted a great deal of attention in network literature.

FIGURE 1 AROUND HERE

The plots suggest that at very high quantiles the distributions all converge to power laws, which distinctively display a straight line on a doubly logarithmic scale. An important characteristic of such distributions is the presence of a fat tail, as hinted by the highly skewness detected. Indeed, extremely high degree values occur more frequently in the real data than for instance in a Gaussian model with the same mean and standard deviation.

In what follows, it is rigorously assessed whether the degree distributions characterizing the labour mobility network behave as power laws. It is worth stressing that, in most practical cases, it is extremely difficult to know for certain whether a given quantity is drawn from a power-law distribution; what can be typically done is rather to verify that the data at hand are consistent with a model in which the given quantity is compatible with a power-law distribution.

The probability distribution of a quantity k is said to be power law if it is drawn from a probability distribution of the form given by the following expression

$$p(k) \propto k^{-\alpha} , \quad (1)$$

where α represents the scaling exponent. Notice that such distribution diverges as k tends to zero, so, provided that $\alpha > 1$, the power-law behaviour must set in only above a certain threshold $k_{\min} > 0$. Moreover, in the case of network degree, the quantity of interest can take only positive integer values, i.e. it

is discrete. Considering these two elements, the expression for the complementary cumulative distribution reads

$$P(k) = \frac{\zeta(\alpha, k)}{\zeta(\alpha, k_{\min})}, \quad (2)$$

where the function ζ is the generalized, or Hurwitz zeta function. The goal of the analysis is now to fit expression (2) to the empirical data; this is done by closely following the procedure proposed by Clauset et al. (2009). The methodology runs as follows: first, a power law of the form given in (2) is fitted to the empirical data using maximum likelihood, simultaneously estimating the scaling parameter α and the lower bound of the scaling region $k_{\min}$; then, the goodness of the power-law fit is evaluated by calculating an appropriate p -value.

TABLE 4 AROUND HERE

Table 4 shows the key results of the power-law fitting. Notice that, in this context, p -values are used to rule out the power-law hypothesis; hence, for the power law to be a plausible model for the data, the p -value has to be high, and vice versa; a relatively conservative choice is to reject the power law if the p -value is less than 0.1. For the in-degree and out-degree, the scaling parameters equal 2.625 and 2.710, respectively; the p -values corresponding to such fits are 0.659 and 0.278, revealing that the power law can actually be a plausible model in both cases. As for the total-degree, the scaling parameter equals 2.633; the p -value corresponding to such fit is 0.137, high enough to accept the power law as a reliable model.

The above results can be summarized as follows: the distribution of links in the labour mobility network is very unequal and it exhibits a fat tail that is well approximated by a power law. This means that in the small world of labour mobility, there are big hubs in a dominant position. By means of long-range connections, hubs tie up many small clusters of firms into a single integrated small-world network.

To quantitatively appreciate the role of such connectors, the analysis now proceeds examining how hub failures affect the connectivity of the SCC. According to the procedure proposed by Albert et al. (2000), imagine that a fraction of employers is hit by a crisis, pushing them to close down the firms, and let then observe how the coverage of the SCC varies in response to such failures. Two series of simulations are performed, one in which increasing fractions of vertices chosen completely at random are deleted from the network (random attack), the other in which the highest degree vertices are progressively removed (hub attack). The results are shown in Figure 2.

FIGURE 2 AROUND HERE

Hub attack proves to have a disruptive impact on the coverage of the SCC, and, consequently, on the small-world connectivity; whereas the failure of an equivalent number of firms chosen at random leads to only marginal changes in the size of the giant strongly connected component. In particular, the removal of just 1% of hubs reduces the SCC by 23%, pushing it from 46.2% to 35.5% of network coverage; on the contrary, the deletion of 1% of vertices chosen at random reduces the SCC by less than 1%. The elimination of 5% of the highest degree vertices cuts the SCC connectivity by more than

a half, down to 22%; and a 10% removal leads to a complete break down of the component, down to 4.2% of network extension. On the other hand, a 10% random removal of vertices reduces the SCC connectivity by less than 6%, leaving almost intact the small-world character of the network.

The explanation for such results is rooted in the Paretian nature of the degree distribution. Since most vertices have only a few connections, and therefore they lie on a very few paths between other vertices, it is unlikely that their (random) removal affects connectivity substantially. But when the removal is deliberately targeted at hubs, each individual failure has disruptive effects, up to deprive the network of its small-world character, via the disintegration of the giant strongly connected component.

These last results lead to the conclusion that: hubs are critical in keeping connected different parts of the network, and such feature is precisely mirrored in the high vulnerability of the system to hub closures. Paraphrasing a terminology risen to fame in the context value chain studies, hubs appear to play a role under many respects analogous to system integrator firms.

7. The main hubs

Making use of the information available in the VWH database, we classify the first 50 hubs according to the sector of activity, and for each sectoral group we display the average workforce, together with the main average network statistics; the results are reported in Table 5.

At a glance, we notice that hubs are mostly large firms. This result is not surprising, it generally true that a larger workforce generates a larger turnover at the firm level that

might result in more network connections either on entry and exit. In the network under scrutiny, the Pearson correlation coefficient between average firm size and total-degree equals 0.57, highly significant, indicating a correlation of moderate intensity.

TABLE 5 AROUND HERE

Examining the data thoroughly, three broad categories of businesses can be detected:

1. long-tradition manufacturing firms;
2. companies involved in services and commerce, typically organized into chains of stores;
3. companies specialized in providing workforce and logistics to third parties.

Among the manufacturing hubs, it is easy to recognize several distinctive activities of the Veneto economy. Four companies produce household and professional appliances, either as final makers or as suppliers of dedicated parts; they belong to the so called ‘Inox Valley’ district, located in the province of Treviso, and they hold prominent positions in their respective market segments at the international level. Two hubs are the undisputed worldwide leaders of the eyeglass industry; they are the major players around which the eyewear district of the Belluno province is organized. Another company is a historical producer of wool fabrics, now mostly involved in the garment sector, which during various decades has fed a vast textile district in the province of Vicenza.

Each of these firms is deeply rooted in a local socio-economic context, and tightly nested into a local cluster of

firms; at the same time, each player has connections spanning different distant areas of the regional and also national economy, indeed, all manufacturing hubs are well-known, historical trade names whose popularity can attract applicant workers from afar.¹⁰

Firms involved in the service sector – comprising hotel trade and tourism, personal and social services, commerce, and catering – share a major trait: they typically carry out business through multiple local units. The ten department stores or supermarkets listed in Table 5 own chains of stores, or point of sales in different cities and towns all over Veneto. The organizations dealing with personal and social services or catering have a highly decentralized structure as well; in this case operations are performed, not only by means of distinct local units, but also through posting workers directly to costumers' places.

A different condition is that of companies specialized in providing workforce to third parties. Two different categories of such firms can be identified: firms providing logistic services as subcontractors, and labour market intermediaries providing temporary workforce. Especially for the latter firms, acting as hubs in the labour market is somewhat inborn in their statutory activity. In Italy, logistic subcontractors represent a long-standing phenomenon, that has been intensifying over time, hand in hand with the increasing segmentation of production processes, and the rising importance of logistics in manufacturing value chains. In contrast, private firms dealing with temporary workers in a range of different fields represent a novelty.

¹⁰ Notice that, not all the major industrial districts of Veneto have a hub figuring among the largest ones. For instance, the goldsmith district of Vicenza, as well as the footwear district of Riviera del Brenta do not compare in our shortlist.

The rapid settlement of a number of private labour market intermediaries – suddenly emerged in the role of hubs – is the result of the policies adopted in the second half of the 1990s, aimed at rendering the labour market more flexible.¹¹ Employment agencies – often local branches of companies operating worldwide – act in a natural way as transmission belts for labour reallocation. Moreover, they treat a variety of occupational profiles, and allocate skilled personnel; whereas, companies specialized in logistic services deal mostly with low-skilled personnel, often recruited within the weakest components of the workforce, as for instance immigrants.

8. Concluding remarks

This essay studies the phenomenon of worker reallocation in the Veneto region by means of a network approach which is new to regional labour market studies.

The empirical investigation is articulated in two parts. In the first one, the paper attempts to assess to what extent the regional labour market is integrated and permeable to worker flows, and it quantifies such properties by comparing the actual network with a small-world one, showing that the inter-firm network examined exhibits a clear-cut small-world pattern. In the second part, the paper asks whether there exist employers playing a strategic role in order to guarantee network connectivity; this question is framed in precise quantitative terms by examining the shape of the link distribution, showing the presence of dominant firms playing the role of hubs for labour flows.

¹¹ In particular, the so called ‘Treu Reform’ of the labour market passed in 1997 significantly increased the scope for workforce intermediation and introduced new forms of fixed-term contracts.

The revealed small-world architecture points to a labour market structure that is permeable to worker flows, and that appears to be capable of effectively absorbing localized cyclical shocks; this is a highly desirable characteristic policy makers might want to preserve and promote. But such connectivity pattern turns out to depend crucially on the presence of a small number of highly connected hubs that span the network from side to side, bridging together distinct local clusters of firms, and keeping the network globally integrated. This latter feature makes the network vulnerable; indeed, the failure of even a small number of hubs most likely results in the split of the labour market into different separated pieces; a more segregated labour market is in turn especially exposed to the harmful effects of asymmetric shocks, which could not be effectively re-absorbed, and hence may cause persistent unemployment in the areas of the economy that remain isolated.

Examining the assembled evidence altogether, some appealing policy implication can be derived; namely, policy makers might want to especially support those firms that guarantee major connectivity to the system, and in this case an effective reallocation policy would hinge critically upon the ability to localize hubs and their connection patterns. Among the most important hubs in Veneto, a few large and long-tradition manufacturing firms are detected. Safeguarding such firms, which are nowadays particularly exposed to the competitive pressure stemming from low-wages countries, could be a first-order goal for policy. The identification and characterization of hubs might offer a criterion for assigning subsidies – as well as extraordinary funds aimed at supporting employment – which are in the discretionary disposal of the Regional Government or other Public Administrations.

Furthermore, the paper shows that temporary staffing agencies, a new family of actors in the regional economy under scrutiny, enter the reallocation market and rapidly rise to hub positions, calling for further investigation.

More in general, a network analysis of labour mobility of the type proposed in this essay seems to offer a suitable toolkit to be put aside more traditional econometrics in order to achieve an effective appreciation and monitoring of structural traits of the reallocation market. Noticeably, the study also reveals that, even in a regional economy whose development has been based for decades on very small organizational dimensions, a few large organizations still play a central role in keeping integrated different parts of the system and allowing for workers to find reallocation opportunities that otherwise would probably be more difficult or impossible to be seized.

The present essay offers a set of validation tests of complex network methodologies for studying labour markets in the context of regional economies; and it provides a number of results with potentially relevant implications for policy. The analysis performed is just a first and incomplete draft of what can still be done by way of applying more systematically network analysis to micro data from employer-employee datasets. This paper is indeed prompted by the great deal of attention network approaches have recently received; for instance Ter Wal and Boschma (2009) claim network analysis has a huge potential to enrich the literature on clusters, regional innovation systems and knowledge spillovers; more in general, in the last couple of years a number of scholars has called for further experimenting network techniques on several new economic issues (Brandt et al., 2009; Schweitzer et al., 2009), recognizing the potential of this technique. The stakes seem to be high and so it is worth investing in this direction.

The present essay lies precisely in this wake, illustrating a possible application of network methodologies to labour mobility on the basis of large-scale linked employer-employee data from administrative source.

Several directions for future developments are in sight. For instance, dynamic network analysis can be performed in order to map the evolution of regional labour markets, again providing both effective monitoring in support of targeted policies and rich information about industrial dynamics; network analysis can also be applied selectively to flows of skilled workers, so as to obtain a map of potential knowledge flows between organizations; or positional indicators of employers in the labour mobility network can be matched with performance data, aiming at disentangling the relationship between these two measures. In order to accomplish a research agenda of this kind, more detailed information about individual professional qualifications and firm performance should be mined, and more sophisticated empirical strategies should be designed in order to properly investigate dynamic networks. This is an exciting challenge for future research.

Appendix: Network definitions

This Appendix provides the formal definitions of the graph-theoretic concepts used in this study (Boccaletti et al., 2006). Let $V = \{i: 1, 2, \dots, n\}$ be a finite set of firms, representing network vertices. For each ordered pair of firms (i, j) , with $i, j \in V$ and $i \neq j$, let $l_{ij} \in \{0, 1\}$ be a link pointing from i to j , with $l_{ij} = 1$ if a flow of workers has passed from firm i to firm j (active link), and $l_{ij} = 0$ otherwise (inactive link); let then $L = \{l_{ij}\}$ be the collection of such links. The set of firms and the set of links form the binary, directed labour mobility network $G(V, L)$, of which an instructive graphical example is given in Figure A1. The total number of vertices in a graph is n , the number of active links is $m = \sum_{i \in V} \sum_{j \in V} l_{ij}$; the number of active links divided by the total number of links gives the network *density*, denoted by $\delta(G) = m/n(n-1)$. The word links is used in the text to refer to just active links.

FIGURE A1 AROUND HERE

The number of links pointing towards i is defined *in-degree* of vertex i , and it is denoted by k_i^{in} ; similarly, the number of links originating from i is defined *out-degree* of vertex i , and it is denoted by k_i^{out} . The *total-degree* of vertex i , indicated by k_i^{tot} , is the sum of the in-degree and the out-degree. In formal terms the following expressions can be written

$$k_i^{in} = \sum_{j \in V} l_{ji} , \quad (A1a)$$

$$k_i^{out} = \sum_{j \in V} l_{ij} \quad , \quad (A1b)$$

$$k_i^{tot} = k_i^{in} + k_i^{out} \quad . \quad (A1c)$$

For the sake of simplicity, in the text the word degree is used to refer to total-degree. The average degree of a network is equal to the average degree of its vertices, denoted by $k(G)$. The vertices with highest degree are usually termed *hubs*. If we think of the degree of a vertex as a realization of a random variable K , the *degree distribution* is then the probability distribution of K , i.e. the probability that a vertex has degree exactly equal to k , and it is indicated by $p(k) = \Pr(K=k)$. In directed networks there exist three different degree distributions for the in-degree, the out-degree, and the total-degree. The *complementary cumulative degree distribution* (CCDD) is denoted by $P(k)$, and it is defined to be $P(k) = \Pr(K \geq k)$.

A *path* from vertex i to vertex j is said to exist either if $l_{ij}=1$, or if there is a set of distinct intermediate vertices $j_1, j_2, \dots, j_n$ such that $l_{ij_1}=l_{j_1j_2}=\dots=l_{j_nj}=1$. A network *component* is a set of vertices which are all reachable through paths, either mutually reachable, obtaining a *strongly connected component*, or just one-way reachable, obtaining a *weakly connected component*. A network may consist of several components, which can be ordered according to their size, i.e. the number of vertices they comprise. A network is said to exhibit a *giant component*, when the largest weakly connected component covers at least 50% of vertices ($n_{wcc} \geq n/2$), the largest strongly connected component covers at least 25% of vertices ($n_{scc} \geq n/4$), and the other components are small (typically of order $\ln(n)$). Giant weakly/strongly connected components are

referred to with the acronyms WCC and SCC, respectively. Path and components are exemplified in the Figure A2.

FIGURE A2 AROUND HERE

The length of a path from i to j is equal to the number of links one has to run along to reach j starting from i . The shortest path from i to j is called *geodesic*, and its length is denoted by d_{ij} . The *average path length* (APL) of a network is defined to be the average length of the geodesics between all possible pairs of vertices in the SCC, and it is denoted by $d(G)$, yielding

$$d(G) = \frac{\sum_{i \in SCC} \sum_{j \in SCC} d_{ij}}{n_{SCC}(n_{SCC} - 1)} . \quad (A2)$$

The set of vertices with which vertex i is directly connected, both on entry and exit, is called (*nearest*) *neighbourhood* of i , and it is defined as $N_i = \{j \in V : l_{ij} = 1 \vee l_{ji} = 1\}$; the number of neighbour vertices of i is thus $\eta_i = |N_i|$. This notion leads to the definition of a metric called *clustering coefficient*. The clustering coefficient of vertex i , denoted by C_i , measures the extent to which the neighbour vertices of i are linked together, forming a densely connected group. Following Watts and Strogatz (1998), the clustering coefficient of vertex i is defined as the ratio between the actual number of links between the neighbours of i , and the maximum possible number of such links. Denoting by u and v two generic neighbours of i , the following expression is obtained

$$C_i = \frac{\sum_{u \in N_i} \sum_{v \in N_i} l_{uv}}{\eta_i(\eta_i - 1)} , \quad (\text{A3})$$

which takes values in the interval $[0,1]$. Vertices with $\eta_i = 1$ are assigned $C_i = 0$. The average clustering coefficient of a network is indicated by $C(G)$, and it is referred to with the acronym ACC. In Figure A3 different graphical examples of clustering for the green vertices are given.

FIGURE A3 AROUND HERE

References

- Abowd J.M. and Kramarz F. (1999) The analysis of labor markets using matched employer-employee data, in Ashenfelter O. and Card D. (Eds) *Handbook of Labor Economics Volume 3*, North Holland, San Diego and Amsterdam.
- Acemoglu D. (2009) The Crisis of 2008: Structural Lessons for and from Economics, CEPR Policy Insight n. 28.
- Albert R., Jeong H. and Barabasi A.-L. (2000) Error and attack tolerance in complex networks, *Nature*, 406, 378-382.
- Amaral L.A.N., Scala A., Barthélemy M. and Stanley H.E. (2000) Classes of small-world networks, *PNAS*, 97(21), 11149-11152.
- Almeida P. and Kogut B. (1999) Localization of knowledge and the mobility of engineers in regional networks, *Management Science*, 45, 905-917.
- Audretsch D.B. and Keilbach M. (2005) The Mobility of Economic Agents as Conduits of Knowledge Spillovers, in Fornahl D., Zellner C. and Audretsch D.B. (Eds) *The Role of Labour Mobility and Informal Networks for Knowledge Transfer*, Springer, Boston.
- Baily M.N, Hulten C. and Campbell D. (1992) Productivity Dynamics in Manufacturing Plants, *Brookings Papers on Economic Activity: Microeconomics*, 187-249.
- Barabasi A.-L. and Albert R. (1999) Emerging of Scaling in Random Networks, *Science*, 286, 509-512.
- Barabasi A.-L. and Albert R. (2002) Statistical mechanics of complex networks, *Review Of Modern Physics*, 74, 47-97.
- Bartelsman E.J. and Dhrymes P.J. (1998) Productivity Dynamics: U.S. Manufacturing Plants, 1972-1986, *Journal of Productivity Analysis*, 9, 5-34.
- Becattini G., Pyke F. and Sengenberger W. (1990) *Industrial Districts and Inter-Firm Cooperation in Italy*, IILS, Geneva.

- Boccaletti S., Latora V., Moreno Y., Chavez M. and Hwang D.-U. (2006) Complex networks: structure and dynamics, *Physics Reports*, 424, 175-308.
- Boschma R., Eriksson R. and Lindgren U. (2009) How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity, *Journal of Economic Geography*, 9, 169-190.
- Brandt A., Hahn C., Krätke S. and Kiese M. (2009) Metropolitan regions in the knowledge economy: network analysis as a strategic information tool, *Tijdschrift voor economische en sociale geografie*, 100(2), 236-249.
- Broder A., Kumar R., Maghoul F., Raghavan P., Rajagopalan S., Stata R., Tomkins A. and Wiener J. (2000) Graph structure in the Web, *Computer Networks*, 33(1-6), 309-320.
- Brown J.D. and Earle J.S. (2008) Understanding the Contributions of Reallocation to Productivity Growth: Lessons from a Comparative Firm-Level Analysis, Upjohn Institute Working Paper No. 08-141.
- Brusco S. (1986) Small Firms and Industrial Districts: The Experience of Italy, in Keeble D. and Wever E. (Eds) *New Firms and Regional Development in Europe*, Croom Helm, London.
- Clauset A., Shalizi C.R. and Newman M.E.J. (2009) Power-law distributions in empirical data, *SIAM Review*, 51, 661-703.
- Cooper D.P. (2001) Innovation and reciprocal externalities: information transmission via job mobility, *Journal of Economic Behavior and Organization*, 45, 403-425.
- Curarrini S., and Feri F. (2006) Reti di trasferimenti di lavoratori, in Regione Autonoma Friuli Venezia Giulia (Ed) *Il mercato del lavoro in Friuli Venezia Giulia – Rapporto 2006*, Trieste.

- Davis S.J. and Haltiwanger J.C. (1999) Gross Job Flows, in Ashenfelter O. and Card D. (Eds) *Handbook of Labor Economics Vol. 3B*, North Holland, San Diego and Amsterdam.
- Davis S.J., Faberman R.J. and Haltiwanger J. (2006) The Flow Approach to Labor Markets: New Data Sources and Micro-Macro Links, *The Journal of Economic Perspectives*, 20(3), 3-26.
- Eurostat, Regional Statistics, <http://epp.eurostat.ec.europa.eu>.
- Fagiolo G., Reyes J. and Schiavo S. (2010) The evolution of the world trade web: a weighted-network analysis, *Journal of Evolutionary Economics*, 20(4), 479-514.
- Fleming L. and Frenken K. (2007) The evolution of inventor networks in the Silicon Valley and Boston regions, *Advances in Complex Systems*, 10, 53-71.
- Fleming L., Mingo S. and Chen D. (2007) Collaborative Brokerage, Generative Creativity, and Creative Success, *Administrative Science Quarterly*, 52, 443-475.
- Foster L., Haltiwanger J.C. and Krizan C.J. (2001) Aggregate Productivity Growth: Lessons from Microeconomic Evidence, in Dean E., Harper M. and C. Hulten (Eds), *New Developments in Productivity Analysis*, University of Chicago Press, Chicago.
- Freeman L.C. (1979) Centrality in social networks conceptual clarification, *Social Networks*, 1(3), 215-239.
- Frenken K. (2000) A complexity approach to innovation networks. The case of the aircraft industry (1909–1997), *Research Policy*, 29(2), 257-272.
- Gabaix X. (2009) Power Laws in Economics and Finance, *Annual Review of Economics*, 1, 255-293.
- Goyal S. (2007) *Connections: an introduction to the economics of networks*, Princeton University Press, Princeton and Oxford.

Goyal S., Van der Leij M.J., and Moraga-Gonzales J.L. (2006) Economics: An Emerging Small World, *Journal of Political Economy*, 114(2), 403-412.

Granovetter M.S. (1973) The Strength of Weak Ties, *American Journal of Sociology*, 78, 1360-1380.

Hamermesh D. (2008) Fun with matched firm-employee data: Progress and road maps, *Labour Economics*, 15(4), 662-672.

Istat, Census 2001, <http://dwcis.istat.it/cis/index.htm>.

Istat, Regional Economic Accounts, http://www.istat.it/dati/db_siti.

Jackson M.O. (2008) *Social and Economic Networks*, Princeton University Press, Princeton and Oxford.

Kauffeld-Monz M. and Fritsch M. (2010) Who Are the Knowledge Brokers in Regional Systems of Innovation? A Multi-Actor Network Analysis, *Regional Studies*, iFirst, 1-17.

Latora V. and Marchiori M. (2001) Efficient Behavior of Small-World Networks, *Physical Review Letters*, 87, 198701.

Milgram S. (1967) The Small World Problem, *Psychology Today*, 1(1), 60-67.

Newman M.E.J. (2000) Models of The Small World, *Journal of Statistical Physics*, 101(3-4), 819-841.

Newman M.E.J. (2003) The structure and function of complex networks, *SIAM Review*, 45, 167-256.

Newman M.E.J. (2005) Power laws, Pareto distributions and Zipf's law, *Contemporary Physics*, 46, 323-351.

Newman M.E.J., Strogatz S.H. and Watts D.J. (2001) Random graphs with arbitrary degree distributions and their applications, *Physical Review E*, 64, 026118.

Newman M.E.J., Barabasi A.-L. and Watts D.J. (2006) *The Structure and Dynamics of Networks*, Princeton University Press, Princeton and Oxford.

- Occari F., Tattara G. and Volpe M. (2001) Gli archivi anagrafici Inps relativi alle imprese e ai lavoratori dipendenti (Appendice 1), in Tattara G. (Ed) *Il piccolo che nasce dal grande. Le molteplici facce dei distretti industriali veneti*, Franco Angeli, Milano.
- Owen J. and Powell W. (2004) Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community, *Organization Science*, 15(1), 5-21.
- Shettkat R. (Ed) (1996) *The Flow Analysis of Labour Markets*, Routledge, London.
- Schweitzer F., Fagiolo G., Sornette D., Vega-Redondo F., Vespignani A. and White D.R. (2009) Economic Networks: The New Challenges, *Science*, 325, 422-425.
- Sorenson O., Rivkin J.W. and Fleming L. (2006) Complexity, networks and knowledge flow, *Research Policy*, 35(7), 994-1017.
- Tattara G. (Ed) (2001) *Il piccolo che nasce dal grande. Le molteplici facce dei distretti industriali veneti*, Franco Angeli, Milano.
- Tattara G. and Anastasia B. (2003) How was that the Veneto region became so rich? Time and causes of a recent success, MPRA Paper No. 18458.
- Tattara G. and Valentini M. (2010) Turnover and Excess Worker Reallocation. The Veneto Labour Market between 1982 and 1996, *Labour*, 24(4), 474-500.
- Ter Wal A.L.J. and R. A. Boschma (2009) Applying social network analysis in economic geography: framing some key analytic issues, *The Annals of Regional Science*, 43(3), 739-756.
- Ter Wal A.L.J. (2010) Cluster Emergence and Network Evolution: A Longitudinal Analysis of the Inventor Network in Sophia-Antipolis, *Regional Studies*, iFirst, 1-18.

- Verspagen B. and Duysters G. (2007) The small worlds of strategic technology alliances, *Technovation*, 24(7), 563-571.
- Watts D.J. and Strogatz S.H. (1998) Collective dynamics of 'small-world' networks, *Nature*, 393: 440-442.
- Watts D.J. (1999) *Small Worlds: The Dynamics of Networks Between Order and Randomness*, Princeton University Press, Princeton.
- Watts D.J. (2003) *Six Degrees: The Science of a Connected Age*, W.W. Norton, New York.
- White D.R. and Houseman M. (2002) The Navigability of Strong Ties: Small Worlds, Tie Strength and Network Topology, *Complexity*, 8(1), 72-81.

Table 1. The four defining properties of small-world networks

	<i>Small world properties</i>		<i>Evaluation criteria</i>
	<i>Description</i>	<i>Network-theoretic definition</i>	
1.	A large number of vertices is reachable from many other vertices in the network	The network exhibits a giant weakly connected component (WCC) and a giant strongly connected component (SCC)	The SCC and the WCC cover at least 25% and 50% of vertices respectively, and the other components are small (typically of order $\ln(n)$, where n is the total number of vertices)
2.	The system is overall parsimonious	The density of links is low	The link density is several orders of magnitude lower than its maximum possible value
3.	Vertices are reachable with little effort	The average path length (APL) in the SCC is appreciably small	The APL in the SCC is of the same or lower order as $\ln(n_{\text{scc}})$
4.	Vertices form cohesive groups	The average clustering coefficient (ACC) is appreciably high	The ACC is at least two orders of magnitude higher than the ACC of a random network of the same size as and with the same in-degree distribution of the actual network

Table 2. Veneto economic statistics

<i>Statistic</i>	<i>Period</i>	<i>Unit of measure</i>	<i>Value</i>
Population	2000	Millions	4.5
Employed persons	2000	Millions	2.1
Unemployment rate	2000	%	3.7
Total GDP	2000	Euros, millions, current prices	112,719
Per capita GDP	2000	Euros, current prices	25,065
Average growth rate of GDP (1991-2000)*	1991-2000	%, constant prices	2.3
Incidence of manufacturing over GDP*	2000	%	32.4
Average firm size**	2001	N. of employees	3
Average firm size in manufacturing**	2001	N. of employees	9

Sources: Eurostat–Regional Statistics; *Istat–Regional Economic Accounts; **Istat–Census 2001

Table 3. Main network statistics

<i>Statistic</i>	<i>Value</i>
N. of vertices	379,391
N. of links	1,899,898
Average total-degree (std. dev.)	10.02 (56.35)
Density	1.32e-05
Size of WCC (as percentage)	373,998 (98.6)
Size of second largest weakly connected component	6
Size of SCC (as percentage)	175,436 (46.2)
Size of second largest strongly connected component	4
Size of IN component (as percentage)	107,336 (28.3)
Size of OUT component (as percentage)	86,398 (22.8)
APL in SCC (as percentage)	4.41
ACC of actual network	3.48e-02
ACC of random network	5.49e-05
ACC of random network based on the actual in-degree distribution	3.53e-04

Table 4. Power-law fit to actual degree distributions

	<i>total-degree</i>	<i>in-degree</i>	<i>out-degree</i>
total n. of observations with $k > 0$	379,391	272,875	292,371
scaling parameter α	2.633	2.625	2.710
cutoff point $k_{\min}$	107	106	56
p -value of the fit	0.137	0.660	0.393
n. of observations with $k \geq k_{\min}$	4,592	1,661	4,298

Table 5. The first 50 hubs by type of activity

<i>Activity</i>	<i>Number of firms</i>	<i>Average workforce</i>	<i>Average degree</i>	<i>Average clustering</i>
Manufacturing	14	1,663	2,416	0.0048
Wholesale	10	1,929	3,326	0.0032
Catering	6	576	2,219	0.0038
Logistic services	6	270	2,966	0.0059
Temporary workforce provision	5	1,013	6,432	0.0023
Personal and social services	5	724	2,725	0.0055
Hotel trade and tourism	2	1,876	7,115	0.0025
Construction	1	388	1,711	0.0013
Advertising	1	165	1,710	0.0081

Fig. 1. Complementary cumulative degree distributions and power-law fits: (a) total-degree; (b) in-degree; (c) out-degree

(c)

Figure 2. Impact of removal of vertices on the SCC

Figure A1. Network with $n=9$, and $m=10$

Figure A2. Strong components (blue, dotted), weak components (orange, dashed), and shortest path from vertex 6 to vertex 4 (green links)

**Figure A3. Clustering coefficient, vertex 1 has: (a) $C_1 = 1$;
(b) $C_1 = 0.4$;
(c) $C_1 = 0$**

(a)

(b)

(c)