Collaboration Networks and Innovation: How to Define Network Boundaries

Galaso, Pablo and Kovářík, Jaromír

Instituto de Economía, Universidad de la República, Dpto. Fundamentos del Análisis Económico, Universidad del País Vasco

8 March 2018

Collaboration Networks and Innovation: How to Define Network Boundaries*

Pablo Galaso†

Jaromír Kovářík‡§

This version: March, 2018

Abstract: Numerous studies in management, sociology, and economics have documented that the architecture of collaboration networks affects the innovation performance of individuals, firms, and regions. Little is known though about whether the association between collaboration patterns and innovation outcomes depends on the network geographical boundaries chosen by the researcher. This issue is crucial for both policy-makers and firms that rely on innovation. This article compares the association between collaboration networks and future patenting between regional and country-level collaboration networks. If we relate future innovation to the global, country-wide network our statistical analysis reproduces the findings of the previous literature. However, we find systematically less important effects of regional innovation patterns on subsequent patenting of innovators. Hence, managers and policy makers should choose the boundaries of the innovation networks that they look at carefully, aiming for integration into larger-scale collaboration communities.

Keywords: innovation, networks, patents, network boundary, boundary specification problem.

* We acknowledge financial support from the *Departamento Vasco de Educación, Política Linguística y Cultura* (IT-783-13), Spanish *Ministerio de Economía y Competitividad* and *Fondo Europeo de Desarrollo Regional* (ECO2015-64467-R MINECO/FEDER and ECO 2015-66027-P) and GACR (17-25222S).

†Instituto de Economía, Facultad de Ciencias Económicas y de Administración, Universidad de la República, Gonzalo Ramírez 1926, 11200, Montevideo, Uruguay.

‡ Dpto. Fundamentos del Análisis Económico & BRiDGE, Universidad del País Vasco UPV/EHU, Av. Lehendakari Aguirre 83, 48015, Bilbao, Spain.

§CERGE-EI, joint workplace of the Charles University and the Academy of Sciences of the Czech Republic, Politckých Vězňů 7, 11121, Prague, Czech Republic.

1. Introduction

Innovation activities usually take place under a framework of cooperative links between agents. Some innovators still operate in an isolated way, but nowadays innovative companies, universities and research centers increasingly share information and resources in cooperative R&D projects, aiming to obtain collective synergies and individual advantages from strategic collaborations (Wuchty *et al.*, 2007; Jones *et al.*, 2008). Previous research in various disciplines has shown that collaboration patterns represent a crucial aspect of innovation processes and determine the success of individual agents and territories (see, e.g. Allen, 1983; Saxenian, 1994; Brusco, 1999). Applying social network analysis, more recent literature characterizes the particular structural properties of innovation networks that enhance or inhibit both present and future R&D outcomes (see e.g. Fleming *et al.* (2007) for analysis of regional performance and Schilling and Phelps (2007) for firm-level analysis; see also Section 2).

Therefore, companies willing to improve their innovation outcomes should take into account the structure of their collaboration patterns while performing research activities with other agents. At the same time, policymakers who seek to encourage innovation should consider how to enhance or inhibit certain features of collaboration networks within and across regions to maximize the knowledge spillovers.

However, innovators typically cooperate simultaneously at different geographical levels, developing projects with other local actors while simultaneously collaborating with players located in distant regions or even in other countries. This poses the question of what network managers or policy makers should look at when designing policies for enhancing innovation. Do regional networks influence innovation outputs in a different way than larger-scale networks? Should managers and policy makers aim for local integration or should they rather enhance integration into national or international research networks?

To address these questions, we first characterize collaboration networks using all the European patents submitted to the Spanish Patent Office between 1978 and 2008. The nodes in our networks are patent applicants and the links represent collaborative relationships among them (Wal and Boschma, 2009;

see below for details). We then break the networks down into the three most important regions in terms of patenting outcomes and apply econometric techniques to contrast how the structure and the position in the full Spanish network compares to the regional networks in affecting future patenting by organizations.

This study is particularly interested in the network boundary problem (Lauman et al., 1989) within the framework of innovation networks. As a result, we follow a particularly strategy. To ensure comparability with previous studies, this research reproduces as far as possible one existing approach modeling the influence of networks on firms' innovation outputs (Ahuja, 2000; Fleming *et al.*, 2007; Schilling and Phelps, 2007). However, we do not—purposefully—follow the literature that employs both local and global network features within one unique estimation framework (Bathlelt et al., 2004; Whittington *et al.*, 2009, or Breschi and Lenzi, 2016). Our strategy thus enables to test directly the boundary specification problem: Is the way in which past collaboration patterns shape future innovations robust to considering regional vs. national innovation networks?

We observe that the associations between the properties of collaboration networks and subsequent patenting detected in the above literature are reproduced using our national network. This makes us confident that any differences between the national and regional networks cannot be attributed to the particular nature of the Spanish innovation market.

The individual, node-level position within the regional and national networks seems to affect future innovations similarly. However, the associations detected using the country-wide Spanish network are not robust to considering the regional collaboration patterns. Consequently, network boundaries seem to play a crucial role in assessing how innovation networks affect future patenting.

Our findings may inform firm-level management policies about what positions to seek in what networks to improve innovation outcomes, and inform policy makers about what aspects of collaboration patterns to promote and at what geographical levels. We also contribute to the understanding of the interaction between geography at different levels and collaboration patterns, a research agenda initiated recently (see Section 2). The classic innovation literature typically assumes local knowledge externalities (e.g. Glaeser *et al.*, 1992), but the network approach to knowledge creation—including the present study—

document important *non-localized* externalities embedded in the overall architecture of collaboration patterns.

The rest of the paper is structured as follows. Section 2 introduces the theoretical framework and proposes our hypotheses. Section 3 describes the data and the methodology of our research. Section 4 presents the results and Section 5 concludes.

2. Innovation, Networks, and Geography

Knowledge creation is commonly analyzed from two perspectives. Economic sociology emphasizes the role of collaborative activities and the resulting structure or network of collaborations (Allen, 1983; Saxenian, 1994; Ahuja, 2000), while economic geography rather stresses the advantages of spatial proximity and clustering as a key determinant of knowledge creation (Jaffe *et al*., 1993). Since inventors and firms are naturally embedded in both social and geographical spaces, the interplay of these two perspectives has more recently received attention (e.g. Lobo and Strumsky, 2008; Breschi and Lissoni, 2009; Whittington *et al*., 2009; Coffano *et al*., 2017).

2.1. Networks and Innovation Performance

On the one hand, research collaborations provide access to relevant information and knowledge, can reduce opportunistic behavior, contribute to the division of risks and cost sharing, help in the division of labor, provide access to a specialized labor force, and may result in a higher market diversification of firms (Katz and Martin, 1997; Nelson and Winter, 1982). On the other hand, such collaborations might have negative influences by increasing the transaction costs and generating lock-in situations (Koput, 1997; Fritsch, 2004). Hence, not all forms of collaborations have to be beneficial.

The structure of research collaborations may be viewed as a network, in which nodes represent innovators (firms or researchers) and a link between two innovators exists if they have collaborated in a research activity. The literature initially used the term *collaboration network* as a metaphor for interdependencies and interactions between firms (see e.g. de Faria *et al*., 2010; Almeida *et al*., 2011). However, recent work presents c*ollaboration networks* as the sets of nodes and links, applying social network analysis to characterize collaboration patterns. In particular, a growing body of research analyzes the impact of collaboration networks on the innovation outputs of organizations and territories (see Pippel (2012) or Phelps *et al*. (2012) for critical reviews).

In this study, we focus on several network features that have received particular attention in the literature from two basic approaches: node-level and network-wide perspective. Node-level approach considers that the relative position of each node in the network can influence its innovation results (Ahuja, 2000; Burt, 2000; Owen-Smith and Powel, 2004; Whittington et al., 2009). In particular, node centrality and connection to large components may facilitate knowledge access, generating positive impacts on their innovation activities. The network-wide approach rather focuses on the patterns describing the whole architecture of interactions among all nodes. From this perspective, several studies found that certain network topologies can impact innovation outputs (Galaso, 2018).

The first network characteristic that we analyze is network closeness, which facilitates information transfers, increasing its accessibility and reliability. Indeed, positive impacts of being close to other network members on innovation have been documented (Fritsch and Kauffeld-Monz, 2010; Fleming et al, 2007).

Another network property widely studied in relation to innovations is the clustering coefficient. Highly clustered networks facilitate the dissemination of tacit, non-codified, complex knowledge, allowing the members of the cluster to pose any potential questions regarding the information shared (Monge *et al.*, 2008; Fritsch and Kauffeld-Monz, 2010; Fleming *et al.*, 2007). In addition, highly connected neighborhoods facilitate trust between nodes and impulse collaborative solutions, as they reduce the possibility of free-riding (Schilling and Phelps, 2007; Uzzi and Spiro, 2005; Cowan and Jonard, 2008).

The small world property – the combination of high levels of clustering with short distances among nodes (Watts and Strogatz, 1999) – combines the advantages of increased clustering and network closeness, because clustering facilitates cooperation and information sharing within clusters and short distances across clusters enhance the transfer of more diverse, non-redundant information across clusters. This hypothesis has been confirmed in some studies (Cowan and Jonard, 2003; Uzzi and Spiro, 2005; Schilling and Phelps, 2007; Uzzi, 2008) but has been shown to be less robust in others (Fleming *et al*., 2007).

Third, high degrees of decentralisation, favored by the existence of different hubs and structural holes in networks, separates non-redundant sources of information, sources that are more additive than overlapping (Burt, 2000).[5] This property may facilitate the creation of diverse ideas within relatively separate parts of networks (Schilling and Phelps, 2007; Monge *et al*., 2008), avoiding collective stalemate (Wal, 2008), and therefore fostering innovation. On the other hand, centralized networks might be efficient structures to setup interoperability among the actors, reducing system disfuctions and enhancing diffusion (Crespo *et al*., 2015).

Lastly, several studies detect a positive impact of the size of the largest connected component on regional performance (Bettencourt *et al.*, 2007; Coffano *et al.*, 2017; Fleming *et al*., 2007). According to these studies, the interconnection of nodes in larger components enhances information flows, knowledge spillovers and cross-disciplinary fertilization, because in networks with smaller and isolated groups new ideas remain unknown outside local contexts, opportunities remain unexploited and, therefore, innovation will be constrained. We thus include a few measures related to the largest component into our analysis.

2.2. Geography and Innovation

The early literature on knowledge spillovers and consequently the growth theories are built on the assumption that knowledge spills over only locally (Glaeser *et al.*, 1992). Extensive literature has indeed

[5]Decentralization and clustering are opposing properties from an individual perspective: if one particular node presents high levels of clustering it cannot, simultaneously, act as a bridge over structural holes, and vice versa (see Phelps *et al*., 2012 on conflicting results related to clustering and structural holes). However, from a global perspective, networks may combine clusters of nodes and high levels of decentralization. Small world networks are a case in point.

reported considerable advantages of geographic agglomeration in many types of economic activities, including knowledge creation (Jaffe *et al.*, 1993; Almeida and Kogut, 1999; Glaeser, 1999; Bettencourt *et al.*, 2007). In certain activities such as transportation or trade, economies of scale apply naturally. As for innovation, physical proximity of firms and industries provides access to a pool of inventors who are embedded in dense network of social relations that span the boundaries of individual firms. Fleming *et al.* (2007) document direct effects of labor mobility on innovation. Moreover, agglomerations of innovative entities are traditionally located close to non-private research institutions such as universities, leading to additional spillovers (Whittington *et al.*, 2009). All these features facilitate localized knowledge flows. Indeed, innovation is closely linked to urban areas (Carlino *et al.*, 2007) and Bettencourt *et al.* (2007) provide evidence that the advantages of such spatial clustering scale up superlinearly with the population size of cities and agglomerations.

2.3. Innovation Networks at Different Geographical Levels

The research on innovation systems has traditionally distinguished between national innovation systems and regional innovation systems (e.g. Carlsson *et al.*, 2002) and therefore pays—despite indirectly—attention to geographic boundaries in innovation and its diffusion. The national innovation system approach analyzes the network of public and private organizations whose activities and interactions initiate, import, modify and diffuse new technologies (Freeman, 1987) and are "either located within or rooted inside the borders of a nation state" (Lundvall, 1992; p. 2). From this perspective, key processes in the creation and diffusion of knowledge take place mainly at a national level (Edquist and Johnson, 1997). The regional innovation system approach rather assumes the existence of relevant differences within countries, placing regions (i.e. geographical areas within countries) at the core of the analysis. According to this literature, spatial concentration and collaboration among actors, as well as institutional and cultural aspects of regions are crucial for knowledge creation and diffusion (Cooke, 1996; Cooke *et al*., 1997).[6] Nevertheless, both perspectives are typically disconnected and built on different assumptions, lack any unifying empirical support, and differ in policy recommendations.

There are only few recent studies that have investigated the interaction between the geography and network perspectives directly. Each of them typically asks one particular question, but their results suggest the role of collaboration networks on innovation cannot be fully understood from a local perspective only. Most of them have taken a region or a city as the unit of observation and asked what determines their innovation outcomes (Bettencourt *et al*., 2007; Lobo and Strumsky, 2008; Breschi and Lenzi, 2016). Bettencourt *et al*. (2007) report that collaboration networks cannot fully explain the superlinear association between innovation and population of U.S. cities. Lobo and Strumsky (2008) directly contrast the relative importance of the non-network features of urban areas against the local structural properties of co-patenting networks, concluding that the former is more important than the latter. Breschi and Lenzi (2016) ask whether short distances are important at the local or global level, or whether they interact. They report that social proximity within regions and to inventors in other cities reinforce each other in the determination of the area-level innovation.

The study most closely related to ours is Whitington *et al*. (2009). First of all, their analysis is at the level of inventors (i.e. individuals and firms). Second, Whittington *et al*. (2009) analyze the interplay between geographical and network closeness, reporting complementary but contingent impacts of both dimensions on firm-level innovation outcomes. They additionally show that the centrality of an organization within both global and local networks has similar effects on its innovation rates. We complement this part of their analysis by studying whether this holds not only for centrality but also other node-level characteristics and whether this is still true for network-level measures.

All this evidence notwithstanding, the existing studies do not answer the main question of our work: May the selection of geographical boundaries of collaboration networks determine the way networks affect collaboration outcomes?

[6] The regional innovation system approach is complementary to other streams of research in regional studies such as innovative milieus (Maillat, 1995), industrial districts (Becattini, 2004), regional clusters (Porter, 2000), or new industrial spaces (Scott and Storper, 1992; Saxenian, 1994).

The discussed literature offers potentially contrasting hypotheses to this question. On the one hand, the purely geographical perspective argues that knowledge mostly spills over locally, suggesting that local collaboration patterns should foster or inhibit firm-level patenting more than applying the country-level networks. Under such view, more distant features of collaboration should exert smaller or even negligible influence in innovation outcomes of firms. Hence, the first hypothesis is:

Hypothesis 1: The effect of network organization and positioning on innovation is stronger for the local, regional networks than for the state-level networks.

On the other hand, the literature analyzing the interaction between geography and collaboration structures indirectly suggests that both local and longer-distance interactions matter for innovation as this research line rarely raises the question of whether this depends on network boundary (Whitington *et al.*, 2009; Breschi and Lenzi, 2016). Hence, an alternative hypothesis may be:

Hypothesis 2: Both the local, regional network organization and the state-level one stimulate innovation equally.

In contrast to the above, there also exist good arguments for a third hypothesis. Since many actors collaborate with both local and more-distant innovators, looking at regional structures only provides an incomplete picture of the market, an issue known as network boundary specification problem in network theory (Laumann et al., 1989). Since it has been shown that non-local collaborations do affect innovation performance (Breschi and Lenzi, 2008; Whittington *et al*., 2009), a purely local analysis of the innovation systems without taking into account such long-distance collaboration may lead to the attenuation biases. That is, using regional collaboration structures might make network properties look as less important than they actually are.[7] The second set of arguments relates to the development of communication and transport technologies that radically lowered the costs of long-distance collaboration (Cairncross, 2001). Indeed, it has been observed that such technologies reduce the geographical barriers in innovation activities (Kaufman *et al.*, 2003). Such a reduction might thus enhance the role of larger-scale collaboration networks, compared to that of the regional structures. As

[7] In fact, since links are not missing at random, there may be even inconsistency issues *caused* by setting the network boundary inappropriately (Chandrasekhar and Lewis, 2016).

a consequence, we propose a hypothesis that is not directly suggested by the existing innovation literature but which actually receives support in our data:

Hypothesis 3: The state-level network organization stimulates innovation more than the local, regional network organization.

The main goal of the following sections is to contrast the differing effect across the two geographical levels.

3. Data and Methodology

To measure innovation and research collaborations, we use all the European patents registered at the Spanish Patent Office from 1978 to 2008. We have constructed a detailed database which enables us to identify the following information on each patent: (1) the date of application; (2) the names of the companies which applied for the patent, which are referred to as innovators throughout the study; (3) the names of the inventors; and (4) the locations of both the innovators and the inventors. We use the European patent applications because these data include the addresses of innovators (applicants) and inventors, while the Spanish patent applications only register the address of the innovators. As explained below, we need both addresses to locate geographically the networks. Regarding the time period, our data covers the period from the starting year of European patent records in Spain to the moment, at which we collected the data.

Previous research using patent data have focused on two different types of collaboration networks, depending on the nodes they consider: networks of inventors (Balconi *et al*., 2004; Fleming *et al*., 2004, 2006) and networks of applicants or innovators (Breshi and Lissoni, 2003; Singh, 2003, 2005; Cantner and Graf, 2006; Graf and Hening, 2008). We follow the second perspective and trace our networks as follows:

- Nodes: The innovators or applicants who register the patent and hold the right to use it. These are mostly private companies, but research institutes, universities, and individuals can also be found among the nodes of our network.

- Links: Following Cantner and Graf (2006) and Graf and Hening (2009), we trace a link between two innovators when they jointly apply for a patent or when the same inventor work for them in different patents. These links are based on the idea that firms or research institutes are related when they work on joint projects or their employees move from one organization to the other.

As in Wal and Boschma (2009) and Graf (2009), we use the addresses of both the innovators and inventors to locate the nodes and links geographically. As a result, some innovators appear in more than one region if either they operate in different regions or if the inventors developing their patents report an address in a different region. We believe that this approach reflects local-level interactions better than disregarding the activities of innovators in multiple regions. See Section 3.4 and Figure 1 for more details.

We use the application date for each patent to determine the timing of collaboration because such date reflects the timing of the collaboration better than the granting date. This information is used to classify the trends in collaboration patterns over time. In our analysis, we split our data into six five-year periods. For each period, we consider only the nodes and links active in the corresponding five years.[8] This is a crucial aspect of our strategy to relate networks to future patenting rates.

To answer our main question, we break the national Spanish network down into regional networks. The geographical boundaries considered to delimit the regions correspond to the NUTS-3 territorial division proposed by the European Union (Eurostat, 2015). We focus on the three largest regions in terms of innovation outcomes: Barcelona, Madrid and Valencia.[9] The objective is to contrast the association between network and patenting at different geographical levels. To that end, we model the link between a set of both node-level and network-level characteristics and the number of patents that companies register in the subsequent period. As mentioned above, the independent variables were carefully selected to reproduce as far as possible the econometric models in the literature (e.g. Ahuja, 2000; Fleming *et al*., 2007; Schilling and Phelps, 2007; Whittington *et al*., 2009). The contribution of this

[8] We choose this time distribution in order to have enough periods with significant data to observe the changes in networks over time. The results are robust to considering different time windows.

[9] It should be pointed out that to compute the network characteristic at national level we use the whole network including the innovators from regions other than the three under scrutiny, but the subsequent econometric analysis only focuses on inventors from the three regions.

analysis is that the network variables are computed from the national as well as the three regional networks and their effects on future patenting are compared.

3.1.Dependent variable

We follow the literature and use the number of registered patents as a measure of novel invention. Patent data present a good indicator of the innovation activities taking place in a given economy (Griliches, 1990; Archibugi, 1992; Andersen, 2001) but also entail several difficulties (see Flemming *et al*. (2007) or Schilling and Phelps (2007) for a discussion). As explained below, some of the difficulties can be addressed, such as patenting differences across time and space (by considering the location of the nodes and periods) and firm-specific differences (by controlling for past patenting).[10]

The idea is to observe how the network architecture influences future innovation results. Hence, to differentiate between current and subsequent outcomes we label our dependent variable *Future Patents*$_{it}$ throughout the analysis. Note that *Future Patents*$_{it}$ is defined as the number of patent applications by organization i in period $t + 1$ and thus corresponds to *Patents*$_{it+1}$. It is a discrete ordinal variable.

3.2.Independent variables

All our explanatory variables refer to network properties.[11] Two of them (*Betweenness* and *In Giant*) measure individual characteristics of nodes based on their positions in the network. Therefore, these individual variables take different values for each actor.

The remaining variables (*Density*, *Centralization*, *Clustering*, *Reach*, *Small World*, *Giant Share* and *Giant Size*) are indicators of the network as a whole, i.e. they measure structural properties reflecting

[10]We cannot control for industry-level differences in patenting, but we hope that past patenting will capture a large part of the variation due to differing patenting cultures in different industries. Since our results corroborate the previous literature that controls for industrial differences, we are confident that they are not driven by different patenting rates across industries.
[11] To save on space, we do not include the formal definitions of the variables. They can be found in e.g. Fleming *et al*. (2007) and Schilling and Phelps (2007).

general patterns of collaboration between all actors. Consequently, these global variables take the same value for each member of one network.

Betweenness centrality

The betweenness centrality of agent *i* measures the number of shortest paths between all pairs of nodes in a network (other than *i*) that pass through *i*.[12] Betweenness is one measure of a firm's access to the information available on the network. However, maintaining links that enable it to remain in the geodesic of a large number of nodes may be costly in terms of time and resources. Consequently, this variable can be either positively or negatively correlated with *Future Patents,* but the literature finds that it generally has a positive impact on future patenting.

In-Giant

The variable labeled as $In\text{-}Giant_{it}$ is a binary variable that takes a value of 1 if node *i* belongs to the giant component of the network and 0 otherwise. It reflects whether a firm is at least indirectly connected to the largest group of innovators. We posit that belonging to this group improves future innovation at the firm as it facilitates access to information and potential sources of collaboration.

Density

Network density measures the number of actual links as a proportion of all the $n(n\text{-}2)/2$ possible connections in a network of *n* nodes. Nodes in denser networks are more interconnected than nodes in less dense networks. Density may increase information dissemination and make it more trustworthy (Fritsch and Kauffeld-Monz, 2010), but high density can lead to the circulation of homogeneous and redundant information (Bettencourt *et al.*, 2007; Lobo and Strumsky, 2008). Consequently, density may both stimulate and inhibit innovation.

[12] We also repeated the whole analysis using the variable *Degree* (instead of *Betweenness*), which measures the number of collaboration partners of each node. The results were almost the same. We focus on betweenness in the final regression because it reflects the position of an agent in the whole network better than the simple local degree.

Centralization

Centralization compares the degree distribution of the network in relation to the most centralized star-shaped network.[13] This property can affect patenting activity in two different ways. First, a centralized network has shorter distances between its nodes, facilitating information flow and stimulating innovation (Schilling and Phelps, 2007) and additionally enhancing coordination mechanisms among the different components of the network (Crespo *et al.*, 2015).[14] However, in highly centralized networks activity is unevenly distributed and too concentrated in the hands of a small proportion of nodes. This provides them with monopoly power over the flow of information and may reduce the aggregate patent production. Therefore, the positive effects of centralization might reach a threshold beyond which higher centralization reduces innovation outputs. Due to the high levels of centralization shown by the Spanish network, we expect this variable to be negatively correlated with the response variable.

Clustering

The clustering coefficient of a network describes the extent to which a network contains highly interconnected groups of nodes. In particular, the average clustering coefficient is the fraction of neighbors of the nodes that are also mutually connected, averaged over the whole population. Again, this can affect innovation in several ways, depending on the levels of network clustering. Under low clustering, an increase in this coefficient may improve innovation results by accelerating the circulation of information. However, when a network is already highly clustered a further increase in clustering might lead to the circulation of homogeneous and redundant information and thus limit positive cross-cluster collaborations (Cowan and Jonard, 2003; Uzzi and Spiro, 2005). Since clustering is generally low in our networks, we expect it to affect our dependent variable positively.

[13] A star-shaped network is an architecture in which one node is connected to all other nodes and all other nodes are connected only to the central node.
[14] Rather than a centralization index, Crespo *et al.* apply the slope of the degree distribution, which essentially measures centralization.

Reach

The *Reach* of a network is the average inverse distance between two nodes in the network (Schilling and Phelps, 2007). This variable reflects how close to each other the network members are. If they are closer, the reach is high, and viceversa. A higher reach may improve innovation outputs by facilitating communication and the flow of information between agents; however, if too high a further increase may lead to homogeneity of information (Uzzi and Spiro, 2005; Uzzi, 2008). The reach is low in our networks, so we expect *Reach* to have a positive impact on our dependent variable.

Small world

Computed as the product of the average clustering and the reach (Schilling and Phelps, 2007), the small-world property represents the degree to which the network can combine high levels of clustering with short distances between nodes. In line with the above arguments, high clustering and high reach may improve the innovation activities in a given network by itself, but the small-world property can show whether the two characteristics reinforce each other. More precisely, clustering may enhance cooperation more when the different clusters are not far from each other (i.e. when reach is high) and the benefits from higher reach might increase when there are close-knit clusters in the network (i.e. when clustering is high); see Cowan and Jonard (2003) and Uzzi and Spiro (2005). Hence, we hypothesize that this variable may be positively associated with future patenting.

Giant Share

Giant Share, the fraction of nodes in the giant component, reflects the relative importance of the main group of innovators a given network. If the majority of agents belong to the giant component, then large-scale dissemination of knowledge is possible (Fleming *et al*., 2007). Consequently, we expect this variable to have a positive influence on the dependent variable.

Giant Size

Rather than the fraction, this variable measures the number of nodes in the giant component. As it rises, the potential for information dissemination and collaboration between its members increases, potentially enhancing innovative activities.

3.3.Control variables

The propensity to patent may of course vary due to non-network firm characteristics or other factors. To account for these possibilities, our modeling approach includes the following control variables:

(Current) Patents

To assess the role of collaboration patterns on future patenting we control for the current innovation outcome of each innovator. The variable *Patents* accounts for a large part of the unobserved heterogeneity in firms' propensity to innovate, such as size, management, industry, etc.

Region Dummies

Different regions naturally provide different innovation and collaboration opportunities. Hence, it is imperative to account for this in the regression analysis. To that end our models include two dummies, one for Madrid and one for Valencia. Their estimated parameters illustrate the different propensities to innovate in these regions, compared to the Barcelona area.

Period Fixed Effects

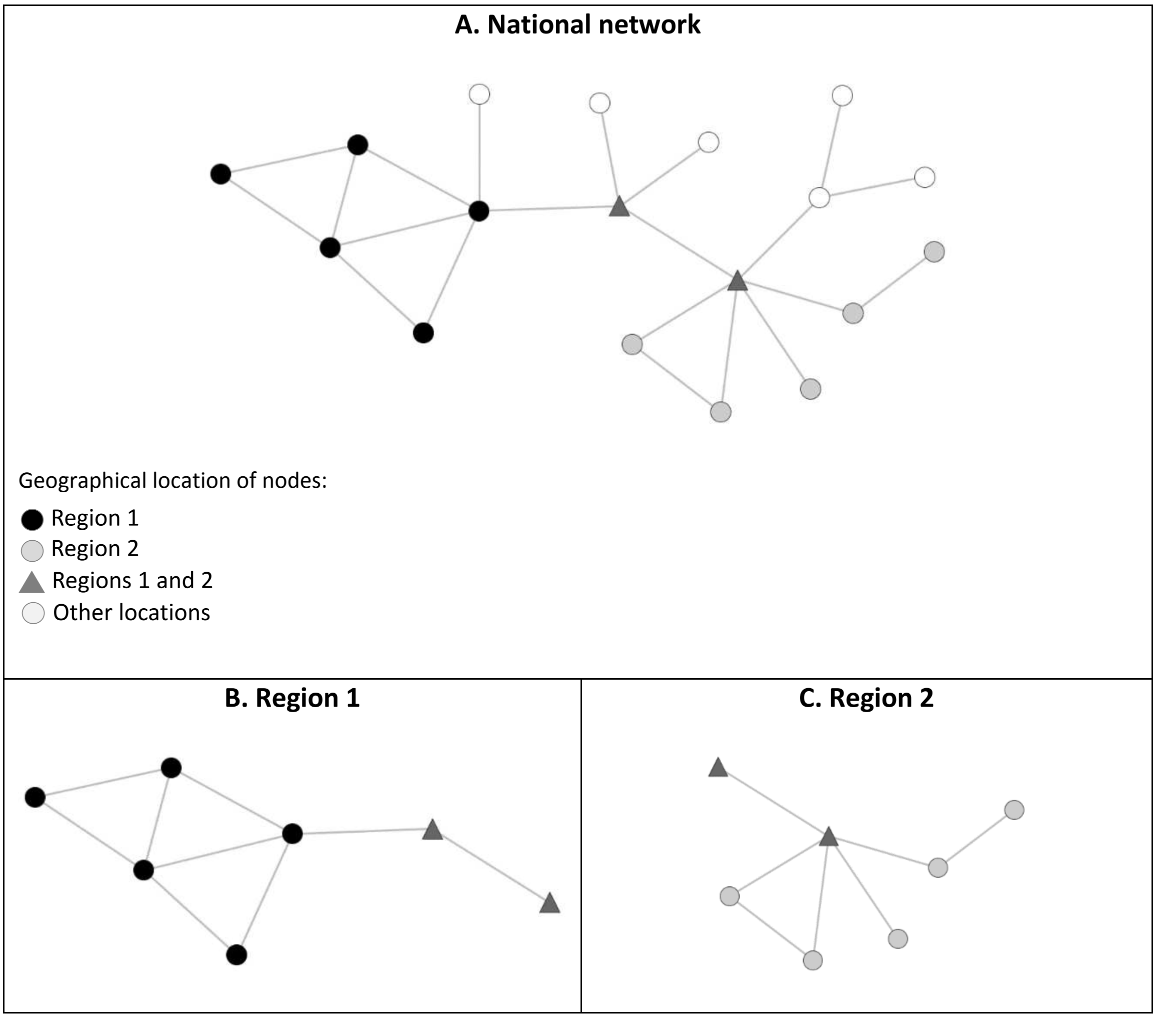
Finally, we control for period fixed effects in our model to reflect potential different propensities to innovate over time. These differences may arise because of macroeconomic, political, or other factors that influence innovation.

3.4.Regional vs. National Boundaries

Given the objective of this study, all network-related variables have been computed using both the regional and national networks. This provides an opportunity to study the influence of regional vs. national collaboration patterns on future innovation outcomes.

Figure 1 illustrates the construction of the regional networks from a hypothetical national innovation structure. Figure 1A represents the national network. The black (grey) circles represent the innovators in Region 1 (2). Hence, these agents are represented in the regional networks in Figures 1B and 1C. The empty circles are innovators who operate in regions that are not analyzed. They are accounted for in the derivation of network measures in the national network, but they are not included in the regional analysis. Naturally, the same would happen to local policy makers or managers, who solely focus on the regional innovation. Lastly, the triangles in Figure 1 correspond to innovators who operate in both Region 1 and Region 2. This may be due to two reasons. Either such innovators have branches in both regions or the innovator and the inventor report different addresses, one in Region 1 and the other in Region 2. We believe that including such innovators provides a more complete picture of the geographical radius of operation of innovators.

Figure 1: The construction of the regional networks from the global national network



Therefore, innovators operating in one of the regions analyzed (or in both of them) may be influenced by the architecture and their position in the national and/or regional networks. Comparing the influence of these networks on innovator's outputs is the aim of our regression analysis.[15]

3.5.Model Specification

Our dependent variable, *Future Patents*, is a count variable that takes integer non-negative values. As a result, a linear regression model is inadequate because it assumes homoskedastic, normally distributed errors, an assumption that is violated in count data. Poisson regression is more appropriate (Hausman *et al*., 1984) but it assumes the equality of the mean and the variance of the variable. Nevertheless, patent data typically present overdispersion and this is also our case. Hence, applying the Poisson model may underestimate the standard errors of coefficients, leading to spuriously high levels of significance (Cameron and Trivedi, 2005). We thus estimate a negative binomial model, which generalizes the Poisson model by incorporating individual, unobserved effects into the conditional mean, allowing for such overdispersion (Hausman *et al*. 1984). This approach is also taken in the related literature analyzing patent data (Fleming *et al*., 2007; Schilling and Phelps, 2007; Whittington *et al*., 2009). We follow these studies closely.

We propose the following general model specification:

Future Patents = *f*(*Betweenness*, *Density*, *Centralization*, *Clustering*, *Reach*, *Small World*, *Giant Share*, *Giant Size*, *In Giant*, *Patents*, *Region*, *Period*)

This general specification omits the indices for innovators, time, and regions, as they differ depending on whether the model is estimated using the national or regional network characteristics.

[15] In case of innovators in multiple regions we estimate, separately, the impact of each regional network on their performance. We assume all of them may be relevant for its innovative activities.

4. Results

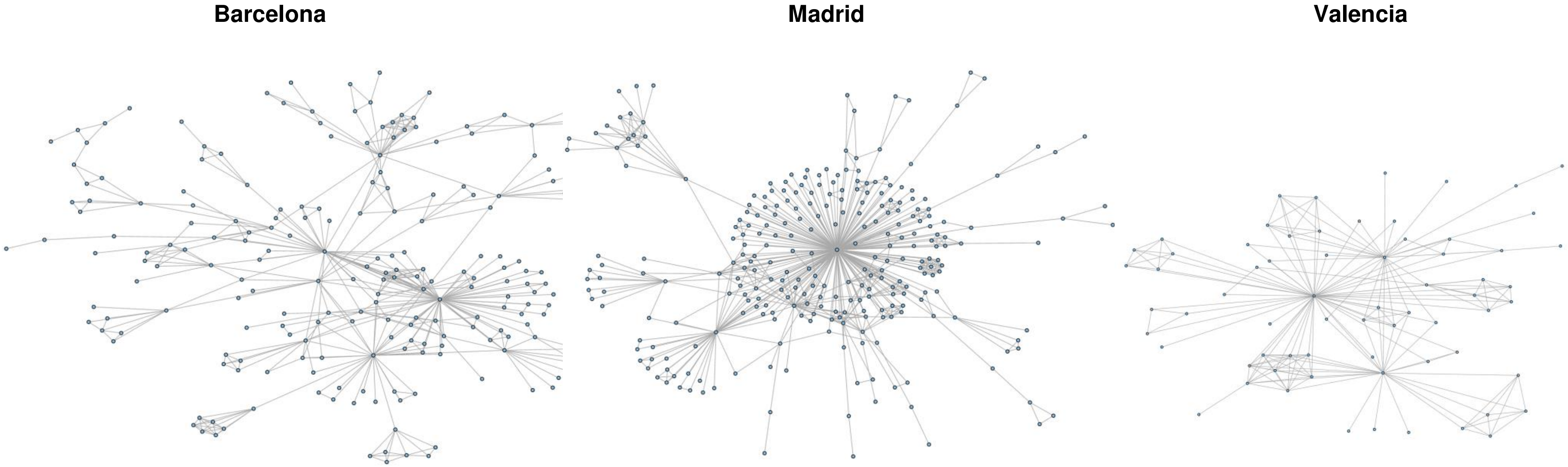
Table 1 presents the general network statistics for our networks. Figure 2 provides a view of the giant component of the Spanish network in the period under scrutiny and Figure 3 illustrates the three regional networks. Almost 60% of innovators (nodes) and collaborations (links) can be found in all three regions under study. Barcelona is the largest regional network in terms of both nodes and links. There are substantial regional differences in the organization of innovation activities. Barcelona has fewer isolated innovators, a lower density, a smaller share of innovators in the giant component, and more inequality in the number of collaborators. In contrast, Valencia and Madrid show larger proportions of isolates, larger shares of nodes in the main component, and more concentration.[16]

[16] The high levels of centralization in Madrid's network can be partially explained by the presence of the Spanish National Research Council (*Consejo Superior de Investigaciones Científicas*), which mainly operates in this region. It is possible to identify this node in Figure 3, as the most central hub, connected to many other nodes.

Figure 2: Giant component of the Spanish Innovation Network (1977-2008).

The nodes are innovators and the links represent both direct and indirect links that took place at some time between 1978 and 2008. The giant components in each period are obviously smaller but they tend to grow significantly over time.

Figure 3: Giant components of the three regional innovation networks: Barcelona, Madrid and Valencia (1977-2008).



The nodes are innovators located in each region and the links represent both direct and indirect links that took place at some time between 1978 and 2008. We observe that there are major differences in how the regional innovation systems are organized.

We estimate several variations of the proposed model and report four of them here.[17] Importantly, we run each regression twice. First, we relate future patenting to the collaboration patterns in the whole country and, second, we repeat each estimation with the regional collaboration structures. Our models account for 385 observations. That is, there were 385 innovators that appeared in any of the networks during at least two consecutive periods.

Table 1: Basic Network Properties in Spain (1977-2008)

		Spain	**Barcelona**	**Madrid**	**Valencia**
Network size	Nodes	8,215	2,459	1,614	604
	Links	5,475	1,558	1,114	458
Density (%)		0.02	0.05	0.09	0.25
Degree	Av.	1.33	1.27	1.38	1.51
	St. Dev.	4.44	2.33	5.25	3.24
Giant component	Size	852	208	278	67
	% of total	*10.37*	*8.46*	*17.22*	*11.09*
Second largest	Size	17	16	10	11
	% of total	*0.21*	*0.65*	*0.62*	*1.82*
Isolates	Number	4,139	1,203	811	300
	% of total	*50.38*	*48.92*	*50.25*	*49.67*
Diameter		11	12	9	5

Table 2 reports the regression results. In what follows, we highlight the main conclusions drawn from our analysis.

First, most network variables influence future innovation in at least one model. Hence, the patterns of collaboration are a major determinant of innovation in our data.

An important question is whether our results are specific to the Spanish innovation system. To that end it is worth pointing out that, using the Spanish-wide network, the estimated effects are in harmony with the literature analyzing U.S. patent data (Fleming *et al.*, 2007; Schilling and Phelps, 2007; Uzzi and Spiro, 2008; Whittington *et al.*, 2009): Centralization is bad for innovation whereas higher clustering and higher reach enhance patenting activities. We detect no effect of the small-world property on

[17] Other variations of the model are available upon request from the authors.

innovation. Even though these findings contrast with e.g. Cowan and Jonard (2003), Uzzi and Spiro (2005) and Schilling and Phelps (2007), they are backed up by studies that do not find a robust effect of the interaction of high clustering and smaller distance (Fleming *et al*., 2007). Hence, we can conclude that our results are not an artifact of the Spanish market for patents.

By contrast, these effects are not generally observed in our analysis using the regional networks. The two node-specific characteristics, *Betweenness* and *In Giant*, matter for innovation and they both matter similarly in the regional and national networks, suggesting that the individual positions of actors in the network are highly relevant for innovation regardless of the geographical boundaries. This corroborates Whittington *et al*. (2009) with respect to centrality. Our analysis shows that their findings seem to extend to other node-level characteristics.

Nevertheless, there are important differences between the regional and national architectures in the effects of the network-level properties on innovation. Overall, the Spanish network plays a larger role in patenting than the regional structures. More precisely, networks which are too centralized lower innovation but this effect disappears for the regional networks if more controls are set in place. Three more variables are found to be significant in models 2 and 4, but only at national level: *Clustering*, *Reach*, and *Density*. The results suggest that local density, measured by the average clustering coefficient, and reach within the country-wide network enhance the creation of cooperative patterns and the circulation of information among actors and higher reach facilitate fast, reliable dissemination. Such effects have previously been detected by Schilling and Phelps (2007) and Uzzi and Spiro (2008). However, these two structural properties are only relevant at the national level in our data. Lastly, network density affects innovation negatively but only if *Clustering* and *Reach* are included in the model (Models 2 and 4) and again only using the Spanish network.

Hence, the role of the collaboration structure in innovation activities is highly sensitive to the geographical boundaries of the network. That is, despite the previous research on innovation networks, we reject both Hypotheses 1 and 2 and rather accept Hypothesis 3.

Table 2: Negative Binomial Regression Models. Dependent variable: Future patents.

	Model 1				Model 2				Model 3				Model 4			
	Regional		National		Regional		National		Regional		National		Regional		National	
	Coef.		Coef.		Coef.		Coef.		Coef.		Coef.		Coef.		Coef.	
Independent Variables																
Betweenness	-0.046	***	-0.001	***	-0.005	***	-0.001	***	-0.004	***	-0.001	***	-0.004	***	-0.001	***
Density	-3.943		-51.225		82.483		-4441.040	*	-24.362		-346.414		10.132		-4410.416	*
Centralization	-7.551	**	0.705		-0.824		-176.684	*	-22.283	*	-103.253	*	-8.672		-176.882	*
Clustering	-		-		2.150		11.840	*	-		-		2.646		11.375	*
Reach	-		-		-117.275		5038.596	*	-		-		-162.649		5017.437	*
Small World	-		-		-		-		-		-		188.504		ommited	
Giant Share	-		-		-		-		0.007		-0.016		15.661		ommited	
Giant Size	-		-		-		-		15.814		117.087		-0.007		ommited	
In Giant	-		-		-		-		0.499	***	0.588	***	0.498	***	0.588	***
Control Variables																
Patents	0.044	***	0.056	***	0.047	***	0.056	***	0.040	***	0.052	***	0.041	***	0.052	***
Madrid	0.377	*	0.063		0.449	*	0.063		0.215		0.020		0.323		0.020	
Valencia	0.078		-0.108		0.047		-0.108		-0.022		-0.085		-0.001		-0.085	
Constant	1.284	**	1.038	***	1.241	*	0.415		0.904		0.679		1.802		0.480	
Model validity																
Log Likelihood	-917.92		-915.94		-916.97		-915.94		-913.09		-910.04		-912.54		-910.04	
Number of obs.	385		385		385		385		385		385		385		385	
Wald chi2	120.55		133.45		125.87		133.45		146.83		177.09		150.06		177.09	
Prob > chi2	0.0000		0.0000		0.0000		0.0000		0.0000		0.0000		0.0000		0.0000	

* p <0.10, **p< 0.05, ***p< 0.01, the models estimated with period fixed effects.

5. Discussion

Previous research reports that the patterns of collaboration in research activities influences future innovation. However, the literature typically sets one particular geographic boundary to characterize the collaboration network in one region, state or country and it is natural to ask whether these findings are robust to considering different geographic boundaries. We argue that this question is of crucial importance for policy makers and business management, especially in knowledge-intensive industries (Owen-Smith and Powell, 2004; Wittington *et al*., 2009). Policy-makers aim to promote innovation by financing public and private research directly as well as indirectly by stimulating collaboration, and this happens at different geographic levels. In a similar vein, firms relying on innovation strategically choose their collaboration partners in their innovation activities and the question is whether they should consider bringing in potential partners in regional, national, and even worldwide collaboration networks and which of them provide the most advantage for their future innovation.

We test these issues using the Spanish patent data from 1978 to 2008, contrasting the role of regional and national collaboration patterns. Our results highlight that the conclusions in the literature may be specific to the geographical boundary that the researcher, policy maker or manager selects for the analysis. Our results corroborate the previous studies mostly applied to U.S. data in many respects (Fleming *et al*., 2007; Schilling and Phelps, 2007; Uzzi and Spiro, 2008; Whittington *et al*., 2009) when the country-wide network is analyzed. Even if no differences are found between the regional and nationwide networks in the influence of the node-level characteristics on innovation, the network-level properties of regional networks show less effect on innovation performance than those of the nationwide networks. Consequently, the association between innovation and network-level collaboration patterns may be sensitive to the geographic area set for analysis.

These findings have important implications. First, they corroborate that networks of collaboration impact future patenting and that this extends to less knowledge-intensive economies such as Spain, since when national collaboration patterns are used the effects detected of the network structural properties on innovation are relatively similar to those found in other studies (Schilling and Phelps, 2007; Whittington *et al*., 2009).

Second, our findings illustrate that scholars should be well aware that focusing on particular regions, such as Silicon Valley or the Boston Area, might mean missing an important part of the general picture of collaboration systems, leading to an overestimation or underestimation of the significance of the collaboration structure.

Third, our results offer an alternative view to the two hypotheses based on the existing evidence. The geographical perspective typically assume that knowledge only spills over locally, while the network approach has abstracted from the role of network boundary and operated as if the boundary did not matter. We show that, in contrast to both views, the boundary matters and the externalities embedded in collaboration networks may be stronger in larger-scale systems than in geographically localized areas. This has important implications for future empirical and theoretical modeling of innovations systems.

Fourth, we inform both policy-makers and managers about what type of innovation patterns and what type of integration they should promote. In particular, they should look beyond their regional innovation systems and aim for integration at higher geographic levels. Our results additionally suggest that country-level and regional policy makers should align their goals regarding research and innovation while promoting collaboration and targeting collaboration patterns.

Needless to say, this study has several limitations. The most important one is the country under study: Spain's economy is less knowledge-intensive economy than that of other countries and more systematic patenting has been observed only recently. This explains why the number of observations is relatively low in the regression analysis, and why only three regions are important in terms of innovation outcomes; it also limits the span of the data. This calls for a robustness analysis using e.g. U.S. data so as to prevent other researchers from drawing too general conclusions from our analysis. However, the internationalization of research activities and globalization lead us to believe that our main conclusion would remain valid in comparisons of e.g. the Spanish vs. the European collaboration network and even in comparisons of the European or the U.S. networks with worldwide collaboration patterns. These hypotheses represent a major question for future research.

6. References

Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. Wiley.

Ahuja, G. (2000). Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study. *Administrative Science Quarterly*, *45*(3), 425–455. https://doi.org/10.2307/2667105

Allen, R. C. (1983). Collective invention. *Journal of Economic Behavior & Organization*, *4*(1), 1–24. https://doi.org/10.1016/0167-2681(83)90023-9

Almeida, P., Hohberger, J., & Parada, P. (2011). Individual scientific collaborations and firm-level innovation. *Industrial and Corporate Change*, *20*(6), 1571–1599. https://doi.org/10.1093/icc/dtr030

Almeida, P., & Kogut, B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science*, *45*(7), 905–917. https://doi.org/10.1287/mnsc.45.7.905

Andersen, B. (2001). *Technological Change and the Evolution of Corporate Innovation: The Structure of Patenting, 1890-1990*. Edward Elgar Publishing.

Archibugi, D. (1992). Patenting as an indicator of technological innovation: a review. *Science and Public Policy*, *19*(6), 357–368. https://doi.org/10.1093/spp/19.6.357

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, *28*(1), 31–56. https://doi.org/10.1191/0309132504ph469oa

Becattini, G. (2004). *Industrial Districts: A New Approach to Industrial Change*. Edward Elgar Publishing.

Bettencourt, L. M. A., Lobo, J., & Strumsky, D. (2007). Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy*, *36*(1), 107–120. https://doi.org/10.1016/j.respol.2006.09.026

Breschi, S., & Lenzi, C. (2016). Co-invention networks and inventive productivity in US cities. *Journal of Urban Economics*, *92*, 66–75. https://doi.org/10.1016/j.jue.2015.12.003

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, 1–30. https://doi.org/10.1093/jeg/lbp008

Brusco, S. (1999). The rules of the game in industrial districts. In A. Grandori (Ed.), *Interfirm Networks: Organization and Industrial Competitiveness*. Routledge.

Burt, R. S. (2000). The Network Structure Of Social Capital. *Research in Organizational Behavior*, *22*, 345–423. https://doi.org/10.1016/S0191-3085(00)22009-1

Cairncross, F. (2001). *The Death of Distance: How the Communications Revolution is Changing Our Lives*. Harvard Business School Press.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban density and the rate of invention. *Journal of Urban Economics*, *61*(3), 389–419. https://doi.org/10.1016/j.jue.2006.08.003

Carlsson, B., Jacobsson, S., Holmén, M., & Rickne, A. (2002). Innovation systems: analytical and methodological issues. *Research policy*, 31(2), 233-245. https://doi.org/10.1016/S0048-7333(01)00138-X

Chandrasekhar, A. G., & Lewis, R. (2016). *Econometrics of sampled networks*.

Coffano, M., Foray, D., & Pezzoni, M. (2017). Does inventor centrality foster regional innovation? The case of the Swiss medical devices sector. *Regional Studies*, *51*(8), 1206–1218. https://doi.org/10.1080/00343404.2017.1329585

Cooke, P. N. (1996). Regional innovation systems: an evolutionary approach. In P. N. Cooke, M. Heidenreich, & H.-J. Braczyk (Eds.), *Regional Innovation Systems: The Role of Governance in a Globalized World*. Psychology Press.

Cooke, P., Uranga, M. G., & Etxebarria, G. (1998). Regional Systems of Innovation: An Evolutionary Perspective. *Environment and Planning A: Economy and Space*, *30*(9), 1563–1584. https://doi.org/10.1068/a301563

Crespo, J., Suire, R., & Vicente, J. (2016). Network structural properties for cluster long-run dynamics: evidence from collaborative R&D networks in the European mobile phone industry. *Industrial and Corporate Change*, *25*(2), 261–282.

https://doi.org/10.1093/icc/dtv032

de Faria, P., Lima, F., & Santos, R. (2010). Cooperation in innovation activities: The importance of partners. *Research Policy*, *39*(8), 1082–1092. https://doi.org/10.1016/j.respol.2010.05.003

Edquist, C., & Johnson, B. H. (1997). Institutions and Organisations in Systems of Innovation. In C. Edquist (Ed.), *Systems of Innovation: Technologies, Institutions, and Organizations*. Pinter.

Fleming, L., King, C., & Juda, A. I. (2007). Small Worlds and Regional Innovation. *Organization Science*, *18*(6), 938–954. https://doi.org/10.1287/orsc.1070.0289

Freeman, C. (1987). *Technology, policy, and economic performance: lessons from Japan*. Pinter Publishers.

Galaso, P. (2018). Network topologies as collective social capital in cities and regions: a critical review of empirical studies. *European Planning Studies*, *26*(3), 571–590. https://doi.org/10.1080/09654313.2017.1406898

Glaeser, E. L. (1999). Learning in Cities. *Journal of Urban Economics*, *46*(2), 254–277. https://doi.org/10.1006/juec.1998.2121

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, *100*(6), 1126–1152.

Graf, H. (2011). Gatekeepers in regional networks of innovators. *Cambridge Journal of Economics*, *35*(1), 173–198. https://doi.org/10.1093/cje/beq001

Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology*, *91*(3), 481–510.

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360–1380. https://doi.org/10.1086/225469

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, *28*, 1661–1707.

Hausman, J., Hall, B. H., & Griliches, Z. (1984). Econometric Models for Count Data with an Application to the Patents-R & D Relationship. *Econometrica*, *52*(4), 909–938. https://doi.org/10.2307/1911191

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic Localization of Knowledge

Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, *108*(3), 577–598. https://doi.org/10.2307/2118401

Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science. *Science*, *322*(5905), 1259–1262. https://doi.org/10.1126/science.1158357

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, *26*(1), 1–18. https://doi.org/10.1016/S0048-7333(96)00917-1

Kaufmann, A., Lehner, P., & Tödtling, F. (2003). Effects of the Internet on the spatial structure of innovation networks. *Information Economics and Policy*, *15*(3), 402–424. https://doi.org/10.1016/S0167-6245(03)00005-2

Koput, K. W. (1997). A Chaotic Model of Innovative Search: Some Answers, Many Questions. *Organization Science*, *8*(5), 528–542. https://doi.org/10.1287/orsc.8.5.528

Laumann, E. O., Marsden, P. V., & Prensky, D. (1989). The boundary specification problem in network analysis. In L. C. Freeman, D. R. White, & A. K. Romney (Eds.), *Research Methods in Social Network Analysis*. Transaction Publishers.

Lobo, J., & Strumsky, D. (2008). Metropolitan patenting, inventor agglomeration and social networks: A tale of two effects. *Journal of Urban Economics*, *63*(3), 871–884. https://doi.org/10.1016/j.jue.2007.07.005

Lundvall, B.-Å. (1992). *National Systems of Innovation: Toward a Theory of Innovation and Interactive Learning*. Anthem Press.

Maillat, D. (1995). Territorial dynamic, innovative milieus and regional policy. *Entrepreneurship & Regional Development*, *7*(2), 157–165. https://doi.org/10.1080/08985629500000010

Nelson, R. R., & Winter, S. G. (1982). *An Evolutionary Theory of Economic Change*. Harvard University Press.

Owen-Smith, J., & Powell, W. W. (2004). Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. *Organization Science*, *15*(1), 5–21. https://doi.org/10.1287/orsc.1030.0054

Phelps, C., Heidl, R., & Wadhwa, A. (7). Knowledge, Networks, and Knowledge Networks A Review and Research Agenda. *Journal of Management*, *38*(4), 1115–1166. https://doi.org/10.1177/0149206311432640

Pippel, G. (2013). The impact of R&D collaboration networks on the performance of firms: a meta-analysis of the evidence. *International Journal of Networking and Virtual Organisations*, *12*(4), 352–373. https://doi.org/10.1504/IJNVO.2013.057282

Porter, M. E. (2000). Location, Competition, and Economic Development: Local Clusters in a Global Economy. *Economic Development Quarterly*, *14*(1), 15–34. https://doi.org/10.1177/089124240001400105

Saxenian, A. (1994). *Regional Advantage*. Harvard University Press.

Schilling, M. A., & Phelps, C. C. (2007). Interfirm Collaboration Networks: The Impact of Large-Scale Network Structure on Firm Innovation. *Management Science*, *53*(7), 1113–1126. https://doi.org/10.1287/mnsc.1060.0624

Scott, A. J., & Storper, M. (1992). Regional development reconsidered. In H. Ernste & V. Meier (Eds.), *Regional development and contemporary industrial response : extending flexible specialisation* (pp. 3–24). London ; New York : Belhaven Press. Retrieved from https://trove.nla.gov.au/version/25122366

Singh, J. (2005). Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science*, *51*(5), 756–770. https://doi.org/10.1287/mnsc.1040.0349

Uzzi, B. (2008). A social network's changing statistical properties and the quality of human innovation. *Journal of Physics A: Mathematical and Theoretical*, *41*(22), 224023. https://doi.org/10.1088/1751-8113/41/22/224023

Uzzi, B., & Spiro, J. (2005). Collaboration and Creativity: The Small World Problem. *American Journal of Sociology*, *111*(2), 447–504. https://doi.org/10.1086/432782

Wal, A. L. J. T. (2013). Cluster Emergence and Network Evolution: A Longitudinal Analysis of the Inventor Network in Sophia-Antipolis. *Regional Studies*, *47*(5), 651–668. https://doi.org/10.1080/00343401003614258

Wal, A. L. J. T., & Boschma, R. A. (2009). Applying social network analysis in economic geography: framing some key analytic issues. *The Annals of Regional Science*, *43*(3), 739–756.

https://doi.org/10.1007/s00168-008-0258-3

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Whittington, K. B., Owen-Smith, J., & Powell, W. W. (2009). Networks, Propinquity, and Innovation in Knowledge-Intensive Industries. *Administrative Science Quarterly*, *54*(1), 90–122.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science*. https://doi.org/10.1126/science.1136099