Research Policy 53 (2024) 104990

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

The emergence and historical evolution of innovation networks: On the factors promoting and hampering patent collaboration in technological lagging economies

Sergio Barbosa[a,*], Patricio Sáiz[b], José L. Zofío[a,c]

[a] Department of Economics, Universidad Autónoma de Madrid, Madrid, Spain
[b] Department of Economic History, Universidad Autónoma de Madrid, Madrid, Spain
[c] Erasmus Research Institute of Management, Rotterdam, the Netherlands

ARTICLE INFO

JEL classification:
N93
N94
K11
C35
C40

Keywords:
Networks
Patents
Collaboration
Innovation
Spain

ABSTRACT

Collaboration and research networks are nowadays central to innovation because they favor knowledge interactions and complex approaches to challenging problems. This study explores the factors underlying the emergence and evolution of innovation networks in the past, using as example the case of Spain, a backward country regarding R&D performance. Combining, for the first time, historical patent data, social network analysis, and discrete choice regression techniques we test distinct institutional, geographical, and sectoral factors that triggered or hampered collaboration over the long term, i.e., the growth in the connections of individual co-patentees within innovation groups. The findings are relevant and demonstrate, inter alia, that in the Spanish case the length of intellectual monopolies did not foster collaboration, while geographical/technological diversification was key to enhance collaborative patterns in the past. The analysis also demonstrates that the likelihood of increasing collaboration over time depended on the initial level of connections (degree) the patentee had, confirming the existence of preferential attachment, even within the context of an emerging and disconnected network. However, belonging to larger innovation groups (size of the network components) did not promote *per se* greater interactions, suggesting that institutional weaknesses and backward innovation trends prevented the existence of positive payoffs from increased connectivity. The results have direct R&D policy implications for both nowadays developing countries and innovation leaders.

1. Introduction

Innovation networks and collaboration structures are drawing increased attention from the academy and R&D managers because nowadays the solutions to complex problems are almost impossible without the integration of distinct sources of knowledge and scientific fields (Owen-Smith and Powell, 2004; Breschi and Lenzi, 2015). Interdisciplinary collaboration favors disruptive advances, new approaches to existing challenges, or ambitious combinations of know-how that several authors call ‘recombinant capital’ (Carnabuci and Operti, 2013; Endres and Harper, 2019). Thus, the way people share and merge ideas, collaborate, or compete when fetching solutions to scientific or technical problems could make the difference between succeeding or failing (Baba et al., 2009; Cantner et al., 2016).

To analyze collaboration, Social Network Analysis (SNA) is a key tool, as it provides an understanding of how agents are connected and how information is shared.[1] Specifically, regarding R&D activities, SNA allows us to analyze the importance of innovation networks by studying their topological properties (see an overview in Phelps et al., 2012; and Pippel, 2013). Specialized literature that studies SNA and collaboration usually refers to patent data as the most suitable proxy to analyze innovation networks. Patents allow both the measurement of the outputs resulting from interfirm collaboration (see, for instance, Ahuja, 2000; Almeida et al., 2011) and the construction of the networks themselves based on distinct ways of patent collaboration (see the literature review in the next section). In the latter case, scholars characterize networks of co-patentees/co-inventors and relate them with economic, entrepreneurial, technological/sectoral, and geographical

* Corresponding author.
E-mail address: sergio.barbosa@uam.es (S. Barbosa).

[1] For an authoritative layman’s introduction to SNA with illustrative examples see Barabási (2003).

https://doi.org/10.1016/j.respol.2024.104990
Received 10 October 2022; Received in revised form 19 February 2024; Accepted 29 February 2024

variables to conclude the features driving the growth of collaboration and its outcomes. In doing so, network topological attributes such as the number of connections individuals have (known as *degree* in SNA), the size of the group to which they belong (*component membership*), or the emergence of large innovation clusters (*giant components*) turn into key variables. In general, studies working with SNA and patents are diverse in the methodology (depending on the focus of the study) but they have one thing in common: they work with recent patent data (from the 1970s onwards) and therefore with mature innovation networks.

In this article, we offer in a novel way a long-term perspective of how cooperation to innovate emerged. Specifically, we delve into the origins and evolution of innovation networks using the case of the Spanish patent system between 1878 and 1939. This was a key stage for the Spanish economic development and industrialization, especially after the First World War (WWI). During this period, international agreements on intellectual property rights (IPRs)—such as the Paris 1883 Convention—were signed and Western economies progressively modernized their patent regimes to defend and promote true invention activity. Spain subscribed to these treaties and guaranteed basic protection, but maintained many characteristics of the first patent systems (patents of introduction, compulsory working clauses, etc.) in an attempt to foster innovation and technology transfer. This remained so during the whole period of our study and even until the Spanish entry into the European Union in 1986 (Sáiz, 2014). The historical data for the study come from the Spanish Patent and Trademark Office (OEPM), where we generated a specific dataset with 68,000 patent applications from residents (and linked non-residents) and with all the patentees disambiguated.

To process this dataset and unveil its characteristics, we rely on SNA methods and develop a historical co-authorship patent network to analyze its topological properties. Specifically, we focus the investigation on the dynamics underlying the growth in the number of connections of the nodes and their component membership, which are vital aspects for any innovation network, because the literature concludes that the larger the nodes' degree and the components' size, the higher the flow of information among actors and throughout the entire network—thereby increasing the likelihood of collaboration in innovation activities (see, for instance, Cantner and Graf, 2006; Schilling and Phelps, 2007). To study the growth of the degree of the nodes we construct the collaboration networks for: 1) a '*Base period*' comprising from 1878 to 1914, i.e., from the Spanish new patent law during the Restoration to WWI; and 2) a '*Cumulative Period*' from 1878 to 1939, i.e., the entire period from 1878 until the end of the Spanish Civil War. First, the notion of a cumulative period traces the patentee's connectivity through time, thereby incorporating the know-how developed through successive patents. As any subsequent patent may draw from preexisting knowledge (links), this is captured by our methodology. Second, we consider only a baseline and a cumulative period because the Spanish innovation network was highly disconnected at the time, with increased connectivity taking place at a slow pace, which resulted in limited network dynamics. Therefore, considering more time windows does not allow us to capture a sufficiently large number of changes to study the determinants of collaboration.

Then, we analyze key transitions from the base to the cumulative period associated with specific increments in the number of connections nodes have (*Degree*), determining the causes that drove collaboration in this emerging network. To study the main factors that facilitated or hampered collaboration among inventors we perform several regression analyses based on the Linear Probability Model (LPM) of the likelihood that patentees increase their degree of collaboration (see Section 4).[2] This allows us to identify the most relevant variables to explain either the emergence and development of complex connections—as Andersson et al. (2019) show for Sweden—or the prevalence of an immature and stagnated innovation network, such as occurred in Spain during the period studied.

For instance, the technological diversification of patentees across industries and their knowledge absorptive capacity, along with their productivity (number of patents), have proven instrumental in triggering collaboration. Besides these, other variables such as the number of assignments or the application for patents of introduction (enabling domestic innovation by transferring—copying—foreign technology) also have a significant effect in achieving higher connected components in the network. In the same fashion, presenting a larger number of connections in the base period increases the probability of widening connectedness in the cumulative period (a property of some networks known as *preferential attachment*).

In contrast, patent duration or registering with a firm status hampered the formation of solid innovation hubs. Unexpectedly, nodes located in larger components of the network in the base period have no advantage in increasing connections in the cumulative period. As the literature highlights (Jackson, 2010, chaps. 6 and 11) collaboration has both benefits (e.g. access to new information) and costs (e.g. time and effort to maintain or extend links). Depending on the institutional and economic frameworks, actors may not have enough incentives to increase collaboration, which seems to occur in the emerging Spanish innovation network during this key period.

In general, our findings contribute to a better understanding of the historical backwardness of the Spanish innovation system and should prove useful in developing policies aimed at fostering social capital regarding R&D dynamics. Certain current key issues—such as the sources of collaboration and cooperation to innovate—deserve also to be analyzed historically and from evolutionary perspectives. For instance, patents of introduction or importation were common in many countries during the nineteenth century to foster industrialization, being progressively banned as innovation systems matured. The Spanish case demonstrates that this institution could facilitate basic cooperation to innovate through technology transfer, which, on the other hand, might delay institutional changes towards the promotion of true domestic invention activity. In turn, this could increase the costs of cooperation to research and the formation of larger structures in the innovation network. Moreover, a central and serious policy implication is that patent length, i.e. the strength of the intellectual monopolies, did not facilitate collaboration. Assuming that cooperation is a necessary condition for the development of successful innovation systems over the long term, then we provide additional evidence to think about the recommendations of scholars claiming to shorten patent length (Boldrin and Levine, 2009). Thus, this investigation offers new results not addressed by the previous literature on SNA and innovation networks, providing a long-term approach for latent topics in innovation studies. That is the case of the discussions on IPR optimal length, the role of knowledge absorptive capacities and multidisciplinary research, or the effects of spatial specialization and institutional environment on collaboration. Notwithstanding, further research in other emerging and mature innovation networks will be necessary to contrast our insights on the origins and evolution of the Spanish innovation system.

The rest of the article is structured as follows. Section 2 reviews the literature on SNA with patent data; Section 3 describes in detail the Spanish patent system, the statistical sources, and the variables used in the study; Section 4 presents the models, specifications, and the LPM regression methodology applied; Section 5 provides results and discussion; and, finally, Section 6 concludes.

2. Innovation networks, SNA, and patent collaboration: a concise literature review

The dynamics of innovation networks increasingly attract economists, business scholars, and other social scientists interested in

[2] See Drivas (2022, p. 249) and Sáiz and Zofío (2022, pp. 262–4) for recent applications of the LPM to explain the emergence of trademarks' geographical specialization in Europe and Spain, respectively.

innovation studies and, specifically, in the circulation of knowledge, the hybridization of ideas, and the complexity of information flows (see, for instance, Fritsch and Kauffeld-Monz, 2010; Tedeschi et al., 2014). To address and understand these processes several scholars have drawn to SNA methodologies and often patent data to proxy innovation networks and interactions among agents. Scholars rely on the information contained in the patents to construct the networks and analyze their structures and the role of collaboration on innovation performance.[3]

The literature uses SNA and patents with distinct analytical objectives but often converges in the methods to build the networks. The most common way is to study collaboration through co-inventors or co-patentees' networks. As patentees can be geographically located and patents/cites technologically classified (Jaffe et al., 1993), researchers can then combine complementary methodological approaches to explore a wide range of topics. The inventors/patentees are the nodes and the co-authorship within a patent are the links in the network. US patents have been the source of many studies related to the geography of innovation. Several of them weight the interactions between the co-inventor networks and spatial dynamics, such as agglomeration phenomena in cities and metropolitan areas that usually reinforce innovation results (Bettencourt et al., 2007; Lobo and Strumsky, 2008; Breschi and Lenzi, 2016), while others analyze network effects at US regional level or in specific innovative areas, such as Silicon Valley or Boston, highlighting the tension between social and spatial dimensions (Fleming et al., 2007; Fleming and Frenken, 2007). Several authors qualify these findings by introducing patent technological and sectoral analyses in US metropolises to conclude that: i) spatial specialization benefits more from gatekeepers (and indirect ties) than diversified cities (Breschi and Lenzi, 2015), ii) the rise/fall of technological knowledge in cities is related to their existing technological base (Boschma et al., 2015) or iii) specialized areas associate with stronger social networks (both internal and external) rather than diversified ones (van der Wouden and Rigby, 2019).

An analogous analysis was developed for other countries using patent data from national offices and the European Patent Office (EPO). One of the most prolific lines studies R&D cooperation and its results in Germany including regional and technological/sectoral perspectives. In this case, the nodes are patent applicants in the national path and the links are common patents or inventors. Findings show that regional specialization in Germany enhances local cooperation compared to external ties (Cantner and Graf, 2004), that inventors' mobility among regions is a key issue in understanding the networks as well as the role of public research centers (Cantner and Graf, 2006; Graf and Henning, 2009), or that the regional knowledge base, on the one hand, and gatekeepers and external connections, on the other, strongly influence network structures and regional innovation systems (Cantner et al., 2010; Graf, 2011; Graf and Krüger, 2011). More recently, related investigations and similar findings have been developed for regions of other European countries such as Switzerland, Italy, and Spain (respectively Coffano et al., 2017; Innocenti et al., 2020; Galaso and Kovářík, 2021). However, very few studies focus outside the US and European frameworks, i.e. on less developed areas of the world, except newly released studies for Latin America. Using Brazilian patents, de Araújo et al. (2019) depict regional interactions among inventors and show the key impact of external links for less innovative and lagging-behind areas. Likewise, using US patents granted to inventors from Latin American countries, Bianchi et al. (2021a, 2021b, 2023) construct international collaborative networks in the area and demonstrate the presence of sparce and highly fragmented structures, the weaknesses of intraregional links (i.e., among Latin American countries), and the crucial role of interactions with high innovative economies outside the region. All these approaches offer fresh insights with relevant implications for innovation policy in less-developed countries. In our historical analysis of the Spanish innovation network, we also include a set of geographical variables to determine whether regional technological specialization and diversification of innovation activities foster or hinder collaboration.

Additionally, few studies focus on how institutions affect collaboration and innovation networks—although this is a key aspect to understanding networks' structures, resilience, or evolution—and, therefore, there has been little opportunity to reflect on policy recommendations and their consequences. For instance, Balconi et al. (2004) analyze co-inventor networks in Italy to differentiate open science and proprietary frameworks finding better connectivity and centrality for academic inventors (more related to the open science world) than those working in proprietary technologies; Cantner et al. (2016) study German governmental policies in renewable energies and find positive effects of such actions on related co-inventor networks, both on their size and structures; and, finally, Menzel et al. (2017) explore how institutional changes impact tie formation, inquiring into co-inventor relationships in ICTs during the dot-com bubble in one of the largest US industrial poles (the Research Triangle Park) with inconclusive results. Likewise, networks' long-term dynamics are crucial to complete institutional approaches. However, there is a lack of historical studies on innovation networks and patent collaboration with only two recent papers. The first compares patent institutions and emerging co-patentee networks in Spain and Sweden from 1878 to 1914 (Andersson et al., 2019), finding common characteristics—like the networks' general disconnection level—but certain key differences—such as bigger and better structured large components in Sweden as well as more openness to foreign influences. The second paper (albeit without using SNA) analyzes the US co-invention framework from 1836 to 1975, highlighting how collaboration and complex knowledge were related over the long term and how co-invention soared from the 1940s onwards linked to increasing technological (and entrepreneurial) development (van der Wouden, 2020). In this investigation we extend the Spanish patent data used by Andersson et al. (2019) to the 1914–1939 period, expecting a more connected and open network through increasing collaboration during the 1920s, a decade of industrialization and growth in the country. However, as we will see in the next sections, the network remains highly disconnected and closed. The lack of collaboration and cooperation to innovate seems to be a critical and long-term problem in Spain.

Thus, there is a pressing need for further institutional and historical research on innovation networks and a challenging opportunity to bring new findings to existing literature. The analysis of long-term dynamics of collaboration patterns and the role of institutions completes and sheds new light on current discussions regarding the effect of cooperation on innovation processes and systems, which may also have potential policy implications for both innovative leading economies and technologically dependent countries or regions. In fact, many of the current scientific and technological leaders were in the past backward or developing areas that experienced a set of changes triggering and driving collaboration and complexity. In general, increasing connectivity reinforces access to information, pushes creativity, reduces opportunism and free-riding behaviors, and fosters knowledge spillovers, multidisciplinary cross-fertilization, and collaborative solutions to innovate (Ter Wal and Boschma, 2009; Schilling and Phelps, 2007; Fleming et al., 2007; Bettencourt et al., 2007; Ter Wal, 2014; Breschi and Lenzi, 2015; Menzel et al., 2017). Other countries, however, had more problems developing competitive innovation systems and fostering knowledge creation and recombination, remaining in backward positions.

In summary, using Spain as case study, this article is the first to analyze the remote roots of collaboration in an emerging innovation network and test the reasons for the backwardness of innovation systems. The findings contribute to the previous literature by identifying institutional and historical factors that trigger or hamper cooperation, which is relevant to current R&D policy-making not only in developing countries but also in leading economies.

[3] For known drawbacks and advantages of adopting patents as proxies of innovation see Griliches (1990).

3. Emerging patent collaboration networks: topologic analysis and stylized facts

3.1. The Spanish patent system: characteristics, sources, and historical data

The first Spanish patent laws of 1826 and 1878 established a patent regime based on a simple registration system, with neither previous technical exams nor oppositional proceedings. The patent office did not openly publish specifications and drawings and there were no specific patent jurisdictions. From the beginning, foreigners could register their inventions, but Spain passed compulsory working clauses (the obligation to produce the invention within the national borders) and patents of introduction (which allowed the registration of other's inventions if they were not implemented in the country). It was a hybrid system designed to protect original inventors but especially to foster industrialization and technology transfer. Spain was one of the signatories of the Paris Convention for the international protection of industrial property in 1883–1884, the base of the current international patent system. As it is well-known (Penrose, 1951), the negotiations provide general principles regarding national treatment for foreigners and priority rights for registering previous patents; temporal protection in international exhibitions; and security that importing one's own patented objects from abroad would not forfeit IPRs. However, the treaty did no prohibit controversial modalities such as patents of introduction or importation and also recognized the rights to maintain compulsory working clauses (Article 5) (a requirement that only from 1925 onwards could be overcome by granting compulsory licenses).

In fact, although Spain had already incorporated one-year priority rights in 1878 and passed modifications to meet the Paris agreement (for instance, the publication of an official gazette from 1886), it maintained both compulsory working clauses and patents of introduction up to 1986, when it joined the European Union. Likewise, Spain kept on granting patents without previous technical examination and leaving the conflicts to the general jurisdiction until the end of the twentieth century. Moreover, the publication of patents did not include drawings and specifications and just the name of the applicant, place of residence, title of the invention, and date. During the nineteenth century, many other countries allowed patents of importation or compulsory working clauses, even developing discriminatory measures against foreigners (Lerner, 2000, 2005; Khan, 2013; Lehmann-Hasemeyer and Streb, 2020). However, from the Paris Convention onwards, and especially throughout the twentieth century, the most advanced Western economies converged in a modern patent regime that reinforced domestic and international IPRs. This process generalized previous technical and novelty examinations and full recognition of the first and true inventors, abandoning those controversial modalities and clauses. International agreements such as the 1970 Patent Cooperation Treaty, or even the birth of the European Patent in 1973 work in the same direction. The actual anomaly in Spain was to maintain the spirit of the nineteenth-century patent regime until 1986, a signal of the country's weaknesses in domestic R&D.

To construct and analyze the Spanish emerging innovation network from 1878 to 1939 we use a historical database (Sáiz et al., 2008) available at the patent office, OEPM, from where we extract the patents applied for by domestic residents plus their non-resident co-patentees.[4] We use patent applications instead of patents because scholars have recently highlighted that using grants—as the previous literature does—may lead to incomplete social networks and estimation biases, something more critical in countries with previous technical examination (Goossen and Paruchuri, 2022). Historical patent documents are similar to modern ones but they differ both in their format (even with handwritten specifications in many cases) and certain data availability. For instance, for the period under analysis, there is patent-related information such as the patent type (invention/introduction); the dates of application, grant, or termination; the annual payments made to maintain the monopoly; whether the patent was assigned; or the international patent classification (IPC), among other administrative issues. Concerning the patentees, there is the availability of personal data (e.g., gender), juridical status (independent/firm), place of residence, and, to a lesser extent, their professions. However, we cannot differentiate patent owners from inventors because patents were mainly registered by independents at that time—a category including inventors, manufacturers, technicians, investors, and alike. Hence, firms were still scarce among patentees during the nineteenth century, and their increasing presence during the first third of the twentieth century in Spain was mainly due to the arrival of foreign companies and corporations rather than to the participation of resident companies. Besides, firms usually patented alone, incorporating previous collaboration processes among agents that it is not possible to disentangle because patents did not list inventors separately yet.

3.2. Variables in the dataset

Table 1 shows the distribution of patents by the original variables in the dataset, i.e., unprocessed variables directly extracted from the historical database. It consists of 67,758 patent applications registered by

Table 1
Distribution of patents by the original variables in the dataset.

Variables	N	%
Duration	67,758	–
0–5 years	59,250	87.44 %
6–10 years	5499	8.12 %
11–15 years	1425	2.10 %
16–20 years	1584	2.34 %
Assignment	2539	3.75 %
Patents of Introduction	13,162	19.43 %
Sectors	67,758	–
Aeronautics	412	0.61 %
Agriculture / Farming	1706	2.52 %
Arms industry	12,433	1.83 %
Basic Metals	1869	2.76 %
Chemical	3841	5.67 %
Communications	591	0.87 %
Construction	3213	4.74 %
Electricity	2834	4.18 %
Gas / Lighting	1254	1.85 %
Lumber	720	1.06 %
Machinery / Equipment	11,405	16.83 %
Mining / Coal	578	0.85 %
Non-Rail Transportation	2645	3.90 %
Paper / Graphic Arts	2679	3.95 %
Railway	938	1.38 %
Sea Transportation	679	1.00 %
Services	14,318	21.13 %
Textile	9898	14.61 %
Food, Beverages & Tob.	6766	9.99 %
Unkown	171	0.25 %
Patents in which at least one of the patentees is a:		
Firm[a]	11,357	16.76 %
Female[a]	1326	1.96 %
Non-Resident[a]	116	0.17 %

[a] The same patent may be registered by several patentees, some of whom may exhibit these characteristics.
Source: Authors' calculations from Sáiz et al. (2008).

[4] See https://www.ibcnetwork.org/e_research_resource.php?id=3 and http://historico.oepm.es/patentes.php

34,663 resident patentees—plus linked non-residents (0.17 %)[5]—that, as expected, were mainly individuals (firms 16.76 %) and males (females 1.96 %). However, while this last variable is directly related to the commercialization of patent rights, '*Duration*' is related to the degree of monopoly power granted by the patent (and therefore to the capacity of deterring competitors). Hence, owning enduring patents may have a double effect. It is a sign of economic success linked to useful and reliable technologies that may encourage further collaboration, but it also gives monopolistic power over a specific technical field that may act in the opposite direction. Notice that most of the patents terminated within 5 years (87.44 %) while very few lasted >15 years (2,34 %)—and only 1,18 % the maximum of 20 years. Likewise, 3.75 % of the patents were assigned during their lifespan. Spanish historical patents were classified according to subclasses of the International Patent Classification by reading the original specifications and drawings.[6] Thereafter, following the typical demand-driven approach by Schmookler (1966), subclasses were grouped into sectors where inventions may have a commercial impact. The sectoral distribution of Spanish patents registered by residents differs from its international counterpart (Sáiz, 2014) as it was more concentrated in low-tech activities grouped in the '*Services*' sector (21.13 %) and light industries such as '*Textiles*' (14.61 %) or '*Food, Beverages & Tobacco*' (9.99 %) while '*Machinery/Equipment*' reached 16.83 %.

As stated, the patent data built by Sáiz et al. (2008) has been previously used by Andersson et al. (2019) in their comparison of the Swedish and Spanish innovation networks. However, while their study ended in 1914, we expand the timeframe to the end of the Spanish Civil War (1939), enlarging the sample from 25,785 patents before 1914 to 67,758 up to 1939 (in both cases applied for residents plus linked non-residents). Notice that with the addition of new variables to the analysis, such as those derived from the inclusion of the geographical or sectoral dimensions, we lose a few observations that have no data for these variables (for instance, applicants with unknown place of residence or patents with unknown sectoral classification).

3.3. Stylized facts on the Spanish innovation network in terms of collaborative components

As patentees can increase their number of linkages over time, we focus our study on capturing the dynamics (evolution) of collaboration in the network. Thus, we use two periods: a *Base period* (1878–1914) between the new Spanish patent law and WWI; and a *Cumulative Period*, (1878–1939), until the end of the Civil War.[7] This method allows capturing incremental changes in the number of connected nodes (between the base and the cumulative period) and their characteristics. In this respect, the links between co-patentees—established through a common patent—are kept during the entire cumulative period.[8] Using SNA tools, we compute the degree and size of the components—to which individuals belong (membership)—for all patentees in the sample for both periods (counting the maximum number of related co-patentees in each period).[9] In doing so, the first step is to create a square matrix of patent/patentee relationships with a total number of 73,823 rows and columns (73,823 × 73,823 matrix), allowing us to obtain the interactions among innovators. This number of patent/patentee relationships represents the sample for all our model regressions.[10]

We identify the following stylized facts characterizing the topological characteristics of the emerging Spanish innovation network.

- *Indefinite degree distribution:* When studying node connectivity within networks, SNA typically confronts two theoretically opposite degree distributions resulting from alternative link formation rules: normal versus power-law distributions. These distributions are respectively consistent with network growth models following random connectivity (Erdös and Rényi, 1959) and preferential attachment (Barabási and Albert, 1999). This last property establishes that the more connected individuals are, the more likely they are to collaborate with others. Thus, preferential attachment is behind the emergence of flat tail degree distributions characterizing the so-called scale-free networks, where few nodes exhibit many links while the remaining are scarcely connected. Strictly speaking, preferential attachment is the principle where the probability of a new node attaching to an existing node is proportional to the degree of the existing node. In the regression specifications, we control for the existence of preferential attachment by including a covariate corresponding to the initial degree of the patentees in the base period. In doing so, we confirm that preferential attachment was at work in the Spanish innovation network. However, the degree distribution did not evolve into a power law, suggesting that, as commonly observed in large real-world networks (especially in an emerging stage), the actual distribution lies somewhere between the domains of 'pure' random and 'pure' scale-free distributions (see Jackson and Rogers, 2007).[11]
- *A low number of connections and size of components (clusters)*: The revised literature in Section 2 consistently highlights the positive role of network density and complex structures on innovation activity and performance (see, for instance, Breschi and Lenzi, 2016; Cantner and Graf, 2004; Lobo and Strumsky, 2008; van der Wouden and Rigby, 2019). However, the Spanish network remained highly disconnected. In the base period (1878–1914), the collaboration network had 2901 connected patentees, representing 19.84 % of the total nodes for the period (shown in the transition matrix for patentees in Table A.1.2 of Appendix 1). Due to the low average degree, these connections took place in 1249 components. Both connected nodes and components are represented on the left side of Fig. 1. Regarding the whole sample in the cumulative period (1878–1939), the number of connected individuals rose to 7362, representing 21.24 % of the total nodes, while the number of components increased to 3092 (both presented on the right side of Fig. 1). Although the relatively higher density of the cumulative network illustrates both an increase in the number of connections and the size of the components, this growth is rather limited.

[5] As far as we want to analyze only domestic clusters of collaborative inventors, we drop out from the database all non-residents without links with residents.

[6] https://www.wipo.int/classifications/ipc/en/.

[7] As historiography usually does, we choose 1914 as a separating year in our analysis because WWI began, disrupting the socioeconomic activity of Europe. That was also a key issue for the Spanish economy. The neutrality in the war facilitated new businesses and expanded industrialization to intermediate and heavy sectors (driven both by external demand and a process of substitution of importations). This also activated new innovation processes.

[8] This implies that even if a patentee passed away or a firm disappeared, their collaboration activity with contemporaneous or posterior co-patentees is recorded in our dataset in a cumulative way. Therefore, a component in the last year (1939) may include co-patentees registering as early as the first year (1878), incorporating, in this way, previous knowledge or know-how and complying with the usual notion of innovation being a cumulative process.

[9] For computations, we rely on the *igraph* package of *R* (https://igraph.org/r/).

[10] Tabulating the 67,758 patents registered by 34,663 patentees into SNA format results in a matrix of 73,823 (patent/patentee) observations containing all the network relationships for the empirical analysis. The matrix of patent/patentee relations contains one observation for each patentee registering a patent. For example, if a patent is registered by three patentees, it corresponds to three observations in the patent-patentee matrix.

[11] We have tested whether the observed degree follows normal or power-law distributions in the base and/or cumulative periods. For this purpose, we use the 'jb' (Jarque-Bera) and 'pwlaw' commands of Stata (for the latter, see Urzúa, 2020). The results of these tests reject the null hypothesis that the degree distributions followed any of these two distributions in any of the two periods.

Fig. 1. Patent network in Spain. Left: base period (1878–1914); Right: cumulative period (1878–1939).
Source: Authors' calculations.

- Non-development of large components: No matter how disconnected an innovation network is, it is critical to analyze the presence of nodes with high degree connectivity that may result in large-size components, characterizing mature innovation systems. In the base period, the most highly connected patentee in the network had 13 connections, creating also the largest component including 17 individuals. In the cumulative period, the most connected individual had a degree of 36, also belonging to the largest component agglutinating 42 individuals. The latter group evolves from the former with the same individual at the center of this star or '*hub-and-spoke*' element. These numbers are not enough to talk about giant components in the collaboration network in Spain (Watts and Strogatz, 1998). Therefore, collaboration in the network did not spread with enough density to reach more connected and complex elements, the key to the development of stronger linkages and larger or giant components that could foster R&D activities. Fig. 2, isolating the largest network components for both periods, shows this limitation.
- Lack of network openness: SNA literature also highlights that openness to foreign ties is essential for innovation (see Lobo and Strumsky, 2008; Whittington et al., 2009; Breschi and Lenzi, 2015). In Spain, non-resident innovators registering from abroad and linked to resident inventors were very scarce (only 74 of the 34,663 patentees were non-residents). Thus, the network was generally closed to interactions with foreign patentees resulting in the absence of foreign direct and personal flows of know-how into the country. This illustrates the isolation of Spain concerning foreign collaboration and implies a loss of innovation opportunities and recombination possibilities with local ideas, directly influencing the development of the collaboration network.

These stylized facts reflect that the evolution of the network, in a period of remarkable economic development, is far from expected. Spain occupied a peripheral technological position in Europe, so determining what were the main characteristics that hampered collaboration is of utmost importance. Likewise, it is also relevant to find out what factors activated connections because cooperation can emerge and become stable over the long term just with a small cluster of individuals who rely on reciprocity (Axelrod, 2006, chap. 1).

4. Modeling the growth of innovation networks by measuring transitions in connections

4.1. Network growth models

Network topologic analysis allows us to extract stylized facts about the Spanish innovation network. Now, to delve into the factors that influenced its evolution, we define a series of transitioning models relating the growth in the degree of the nodes to the characteristics of the patents and patentees observed in the base period:

$$Y^{b,c}_{Model\#} = f\ (Patent/Patentee\ Factors, Network\ Factors, Geographic\ Factors, Sector\ Factors) \quad (1)$$

$Y^{b,c}_{Model\#}$ is a binary variable capturing the growth in the degree of a node from the base period *b* (1878–1914) to the cumulative period *c* (1878–1939) if a *specific threshold* is met. If a patentee reaches the connection threshold established by a particular model ($\# = 1{,}2{,}3$), the dependent variable adopts a unitary value, $Y^{b,c}_{Model\#} = 1$. Otherwise, if the threshold is not met, $Y^{b,c}_{Model\#} = 0$. We study specific thresholds that represent qualitative changes in the configuration and magnitude of connectivity in the network, related to both the increase in the degree of the actors and the emergence of large-size components. These are key drivers in the development of innovation systems as justified below (see Schilling and Phelps, 2007; Bettencourt et al., 2007; Ter Wal, 2014; Jackson, 2010):

1) First, overcoming isolation is a pre-requisite for the emergence of any network (Barabási, 2010, chap. 4; del Fresno, 2015), so we analyze what we term *'Essential change'*, associated with a patentee's transition from being isolated in the base period (*Degree* = 0) to being connected in the cumulative period (*Degree* > 0). In the Spanish innovation network until 1914, <20 % of the patentees were connected in the initial period (see the transition matrix in Table A.1.2 of Appendix 1), and <6 % of the components (groups) were larger than three. Consequently, any substantial expansion of the innovation network requires, first, to overcome this generalized stage of initial isolation. We identify this change as

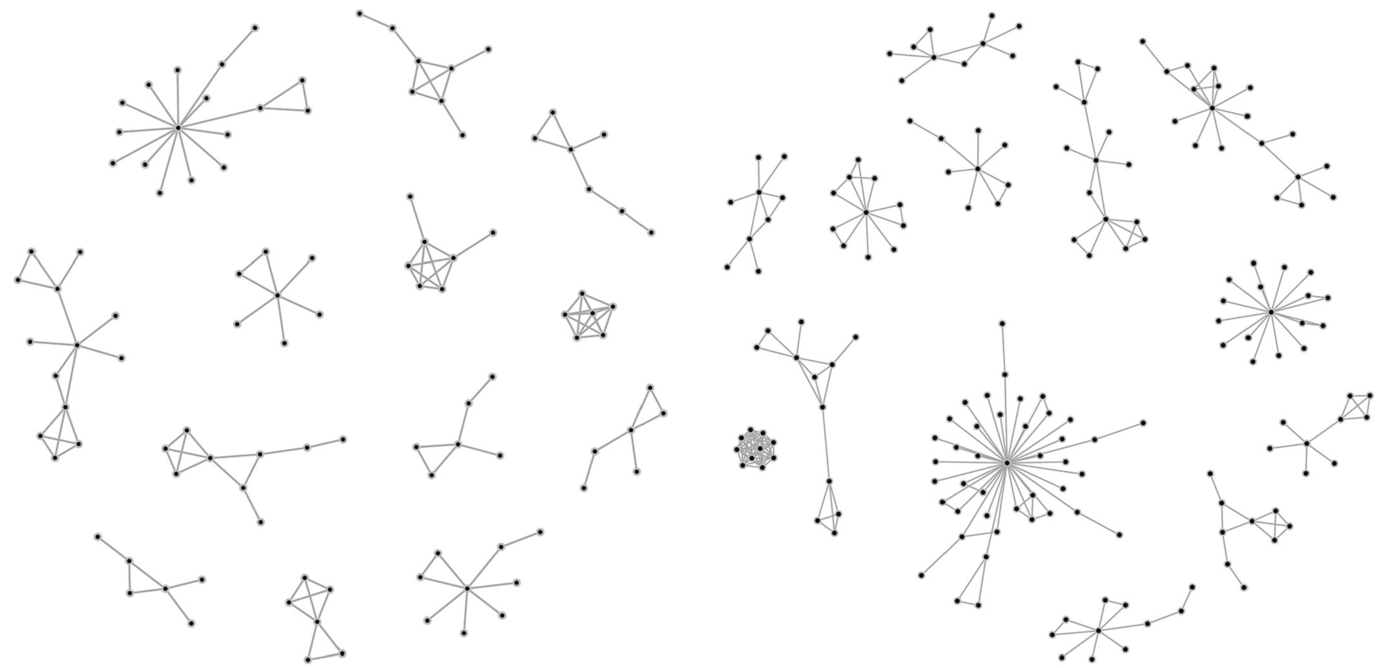

Fig. 2. Network's largest components. Left: base period (1878–1914); Right: cumulative period (1878–1939).
Source: Authors' calculation.

Model 1, $Y^{b,c}_{Model\,1}$: (0; >0). The main incentives for independent patent holders to connect with others were twofold: search for financing and/or technical expertise. During the early industrialization processes the lack of capital was a compelling problem and many individual inventors sought financial support. One way to achieve this was by sharing the patent ownership with businessmen or manufacturers. It was usual to find single patentees associated in further patents with their financial partners. Likewise, as occurs nowadays, inventors could contact and cooperate in seeking technical complementarities to develop new ideas. Thus, there were also patents shared by two or more engineers with distinct backgrounds and expertise. All these kinds of financial or technical alliances could evolve into limited companies or other types of firms. However, corporate patenting was still scarce before the 1920s even in the most industrialized countries (Nicholas, 2010, 2011a), a phenomenon that progressively developed during the second half of the 20th century.[12]

2) Second, we compute the model *'Dyads and Triads'*. This model addresses how paired co-patentees and trios are capable of consolidating collaboration even if within relatively small components. Looking at the transition matrices of the patent/patentee dataset of observations in online Appendix 1 (Tables A.1.1 for degree and A.1.3 for the size of the component), we observe that for already connected individuals (i.e., excluding Model 1), most of the observed transitions correspond to: a) individuals connecting from dyads (components of size 2 where members have a degree of 1) into groups with 3 or 4 individuals (thereby increasing their degree to 2 or 3); and b) triads (either open or closed trios) increasing their degree from 2 to 3. Following the previous notation, we denote the first type of changes as (1; 2, 3), i.e., evolving from a degree of 1 into 2 or 3, and the second type as (2; 3). Therefore, we define a second model that studies these two types of changes aggregately, i.e., Model 2, $Y^{b,c}_{Model\,2}$: (1; 2, 3) ∪ (2; 3). Realizing larger connections, even if to this limited extent, represents both a qualitative and quantitative step forward in collaboration given the limited association levels that existed in the Spanish innovation network. Studying the factors determining these transitions is fundamental as they may trigger further growth resulting in the emergence of larger and more densely connected components. In addition to preferential attachment, we also note that the first type of transition, (1; 2, 3), embeds changes corresponding to the realization of the *principle of triadic closure*, by which two individuals who are not directly connected but are associated with a common third individual, are more likely to connect among themselves than with other people. The transition from an open triad to a closed one naturally favors the emergence of community structures with larger components and fatter-tailed degree distributions (Bianconi et al., 2014; Kali, 2003).[13] Collaboration with "partners of partners" eases long-distance connections and drives network evolution over time (Ter Wal, 2014).

[12] Many examples can illustrate the above statements. For instance, the well-known Spanish civil engineer Jorge Loring Martinez (https://es-m-wikipedia-org.translate.goog/wiki/Jorge_Loring_Mart%C3%ADnez?_x_tr_sl=es&_x_tr_tl=en&_x_tr_hl=es&_x_tr_pto=wapp), who specialized in aeronautics, patented on his own advances in airplanes before and after 1914 (Spanish patents 56,912 of 1913 and 101,047 of 1927). However, between 1918 and 1923 he worked and collaborated with Claudio Baradat Guillé—a mechanical engineer and prolific inventor with expertise in combustion engines—and patented jointly six inventions related to aircraft engines and propellers (Spanish patents 68,049; 70,669; 70,908; 71,433; 71,484; and 85,802). Claudio Baradat himself had individual patents mainly on cinematography, but he needed capital to develop his inventions related to automobile engines, so he associated with a businessman named Federico Esteve Anglada who supported him financially. There are 42 Spanish patents shared by Claudio Baradat and Federico Esteve from 1921 to 1931. In fact, in 1922 a firm called *Cortina, Baradat, and Esteve* was created to manufacture automobiles (https://ca-m-wikipedia-org.translate.goog/wiki/Baradat-Esteve?_x_tr_sl=auto&_x_tr_tl=en&_x_tr_hl=es&_x_tr_pto=wapp), although there are no Spanish patents granted to the company.

[13] In this regard, we have studied transitions from open triads to closed triads. There were 338 open triads in the base period, but none of the two unconnected individuals (or all three members) jointly registered a new patent that would have developed a direct link among them, closing the triad. This issue shows the low connectivity of the network and the poor collaboration dynamics in environments unfavorable to innovation activities.

3) Third, and finally, we want to analyze in more detail the consolidation of larger connectivity levels responsible for the emergence of greater and, eventually, giant components. Thus, we study the transition to *'High connectivity'* consisting of individuals with four or more ties in the cumulative period: Model 3, $Y^{b,c}_{Model\,3}$: (1, 2, 3; >3). High connectivity within large components is important in innovation networks because it brings in unique capabilities, resources, and knowledge that can enable patentees to achieve their goals. A high level of connectivity can provide economies of scale, access to diverse markets, and contact with a wider range of specialized expertise. It can also help to attract new participants and stakeholders, increase the breadth and depth of knowledge sharing, and facilitate the transfer of technology and know-how. In addition, it increases the prospect of raising financial resources to fund R&D activities and product (prototype) and process developments (Fleming and Frenken, 2007; Graf, 2012). Overall, transitioning from a low-density and fragmented innovation network like that existing in Spain to a more densely interconnected network through higher connectivity plays a crucial role in driving progress within innovation networks. As in the case of the previous model we choose this specific threshold based on the values of the transition matrices showing that a significant number of nodes with a degree larger than three emerge in the cumulative period, capturing a relevant number of shifts in connectivity (see, once again, Tables A.1.1 for degree and A.1.3 for the size of the components). In this model, we do not include transitions of isolated individuals with *Degree* zero in the base period because these changes are already studied in Model 1. Since leaving isolation accounts for most of the transitions, including this change would result in a largely overlapping sample, driving the regression results towards those of Model 1. Besides, Model 3 allows us to rigorously test the existence of preferential attachment in the network (a necessary but not sufficient condition for the emergence of power-law distributions). This is accomplished through the inclusion of *Degree* and *Component Size* as explanatory variables given that Model 3 is the only specification that exhibits sufficient variability in both regressors. Indeed, the difference with Model 2 is that we consider all observed connectivity growth above three (>3) (i.e., at least four links). Although the difference may seem small quantitively, it is large when taking into consideration that the number of transitions of patentees starting with a degree of 1 or 2 to a degree of 2 or 3 (Model 2), is the same as those transitioning to a degree larger than 3. That is, while Model 2 studies the consolidation of collaboration among pairs and triads and into quartets, Model 3 focuses on the transition to degrees larger than 3 (that requires components with >4 individuals). Therefore, this model studies the emergence of high-degree individuals (also called *Hubs*), who play a key role in network growth by triggering links and can increase rapidly the average degree of connectivity, resulting in a more developed and densely connected network.[14]

4.2. Models' specifications: Explanatory variables to be included in the transitioning models

The factors (covariates) included in the different models (1) are measured in the base period and grouped in the following specifications considering their incremental inclusion in subsequent regressions:

$$Y^{b,c}_{Model\#} = f\,(Average\,Duration\,of\,Patents\,per\,Patentee\,(ADPP), Assignment, Prolific, Relative\,Technonological\,Diversification\,(RTD), Firm\,Importer, Female, Non-Resident, Degree, Component\,Size, Geographical\,Specialization\,(RGTPS), Geographical\,Diversification\,(RGTPD), Population\,(PatentsPC), Sector\,Dummy). \quad (2)$$

1) Baseline specification: here we include the basic attributes related to the patent and the patentee as well as factors controlling for their characteristics within the network:

- *Average duration of patents per patentee*, '*ADPP*'. Measures the average number of years that all the patents registered by the same patentee are alive, i.e., all observations for the same owner share the same value. This measure aims at capturing both the invention value (associated with economic and social success that may enhance collaboration) and the monopolistic power granted by patents (resulting in extra rents, which may be an incentive to deter further collaborations). Therefore, higher values of ADPP mean that the value of the intellectual monopoly generated with the grant of the patent is higher, and this issue may positively or negatively affect collaboration.
- '*Assignment*'. Measures the number of assignments during the patent life. Provides a proxy for the commercial value of the invention. A patent generates revenues when someone takes a license to it and pays royalties. The number of assignments captures the willingness to pay by third parties for using a patent in their innovation activities, whose utility is revealed by the market value of the assignment.

Combining the above two variables, we can conclude that a patentee with higher ADPP and patent assignments offers lasting inventions of relevant commercial value (Harhoff and Wagner, 2009).[15] Thus, our focus is to observe if owning patents of higher value triggers collaboration.

Other relevant factors facilitating increments in collaboration are the productivity of patentees and the technical breadth of their innovations:

- '*Prolific*'. Measures the number of patents per patentee. The larger the number of patents an individual holds, the greater the likelihood that the patentee will increase his/her collaboration with new individuals in the future. Patentees owning a high number of patents exhibit unique characteristics, summarized in larger innovation experience, and cast positive spillovers (externalities) on other collaborators (Zacchia, 2018). Thereby they are catalysts in the innovation networks. This feature is key to analyzing a network because prolific patentees can be seen as hubs in the innovation network, connecting individuals who may be separated in the product space, while fostering innovations (Granovetter, 1973).
- *Relative Technological Diversification, 'RTD'*. Measures the number of different technological/industrial sectors in which a patentee is present. In an emerging and historical network, we expect that the more diversified an inventor/patentee is, the higher their knowledge *absorptive capacity* and the possibility to create new links (Peters and Johnston, 2009). This variable complements the information in '*Prolific*', i.e., how active are individuals patenting, by capturing the breadth of the innovation across sectors.[16]

[14] We have also run regressions with transitions starting from lower degrees such as (1, 2; >2)—whose results are included in online Appendix 2 as a robustness check of those obtained for Model 3 (see Table A.2.2)—and for higher degrees (1, 2, 3, 4; >4), but in this last case the number of observed transitions reduces sharply, so we cannot draw reliable conclusions (results for this regression are available upon request). Therefore, we select degree >3 to be the relevant threshold.

[15] The interaction between '*Average duration of patents per patentee, ADPP*' and '*Assignment*' is one of the crossed effects considered in an extended regression aiming at providing robustness checks of our regression results, see Tables A.2.2 and A.2.3 in online Appendix 2.

[16] '*Prolific*' and '*Relative Technological Diversification*, *RTD*' are also interacted in our regression with crossed effects, see Tables A.2.2 and A.2.3 in online Appendix 2.

Additionally, we identify the following variables as key characteristics of the patentees:

- '*Firm*'. Dummy variable indicating if the patentee is a firm. Firms can directly implement innovations in the market and exploit the inventions, being able to recombine the existing technologies to create new knowledge and innovations (Carnabuci and Operti, 2013). Firms are crucial to investment activities (Rubens et al., 2011) because knowledge-sharing happens predominantly within organizations due to proximity, frequency of interaction, etc. (Katz and Allen, 1982; Goossen and Paruchuri, 2022).
- '*Importer*'. Dummy variable capturing if patentees rely on the so-called 'patents of introduction'. This type of patent aims at protecting foreign inventions in domestic markets at the expense of the original creators; thus, patentees holding only patents of introduction cannot be considered inventors, because they innovate through technology transfer. This variable is computed as the percentage of patents of introduction over all patents an individual owns.
- '*Female*'. Dummy variable capturing the gender of the patentee. Only 1.96 % of individuals were females (see Table 1), and most likely dependent on their male relatives. This gender dependency may induce co-patents and therefore the illusion of collaboration (Mauleón and Bordons, 2014).
- '*Non-Resident*'. Dummy identifying if the patentee is established abroad. International connections work as major channels or 'global pipelines' that increase knowledge production (Singh et al., 2016). It has been shown that the lack of connectivity to the outside world is a reason to fall into lock-in non-innovative development paths (Gazni et al., 2012; Miguelez et al., 2019)

We now consider the inclusion of two relevant network indicators of the patentees influencing the likelihood of transitioning[17]:

- '*Degree*'. As stated in Section 3.3, in networks driven by preferential attachment the nodes with higher connections are more likely to make new links than those with lower degrees. Thus, this variable controls for this principle, by measuring the effect of the number of connections the patentees had in the base period. We expect a positive sign of *Degree*, meaning that the larger the initial value, the higher the probability of increasing collaboration.
- '*Component Size*'. This variable controls the effect that the size of the components has on connectivity. It counts the number of individuals that belong to the component of each patentee in the base period. In this regard, in a qualified or blended version of the principle of triadic closure and preferential attachment, we expect that individuals belonging to large groups are more likely to increase their connectivity in the innovation network, because: a) slightly connected members within the group already have a pool of commonly related patentees; and b) the larger the number of members in the group, the higher the access to technical information and the likelihood of increased connection.

Although the initial hypothesis for *Degree* and *Component Size* is a positive sign, it is necessary to highlight that link formation brings both opportunities and risks (Jackson, 2010, chaps. 6 and 11). On the one hand, the rise in connectivity and membership in large components yields benefits such as access to a greater pool of information and knowledge, as well as improved positioning within the network. On the other hand, collaboration has costs such as time and effort required to establish and maintain links, which can be determinant depending on the institutional environment and its weaknesses (degree of trust in partners, level of industrial piracy, effectiveness of the judicial system, etc.). Therefore, merely belonging to a larger component and gaining increased access to information may be not enough to enlarge collaboration and could even be a disincentive to increase connectivity. Innovators and patentees could leverage such knowledge to their advantage by taking a more selective approach to collaborations or by extracting and potentially free-riding technological information, etc. Moreover, in certain circumstances, once an agent achieves a certain level of connectivity, further increases may multiply the costs compared to the marginal benefits reversing the sign of these variables from positive to negative. This could help to elucidate specific institutional or economic limits for network connectivity, which in the Spanish emerging innovation network seems to occur with 5–6 connections.

2) *Geographic specification:* We introduce in our model this dimension that has proven key in explaining the growth of innovation systems through geographical proximity and hubs. We compute three indicators that summarize information on the geographical distribution of patents by sectors, along with relative market sizes in the base period (1878–1914):

- *Relative Geographical and Technological Patent Specialization, 'RGTPS'*. Several studies show the positive impact on collaboration of 'domestic (or local) contributions', highlighting the importance of the geographical dimension within a country to develop stronger linkages (Ortega, 2011). Following gravity models, knowledge exchanges intensify when actors are closer geographically (Sáiz and Zofío, 2022). By sectors, we construct an indicator of patent concentration that captures the relative importance of the province where the patentee operates to the national distribution of patents (see Table 2). We expect those patentees operating in provinces with—higher than the national average—sectoral patent concentration, such as Madrid or Barcelona, to be more connected. The indicator is defined as follows:

$$RGTPS^{b}_{s,p} = \frac{Share\,RP^{b}_{s,p}}{Share\,RP^{b}_{s,country}} = \frac{RP^{b}_{s,p} \Big/ \sum_{s=1}^{S} RP^{b}_{s,p}}{\sum_{p=1}^{P} RP^{b}_{p,s} \Big/ \sum_{s=1}^{S}\sum_{p=1}^{P} RP^{b}_{p,s}}, \quad s = 1,\ldots,S, p = 1,\ldots,P. \tag{3}$$

This expression shows the relative patent specialization in sector s of province p in the base period b, defined as the share of *registered patents* (*RP*) in that industry, location, and time, $RP^{b}_{s,p}$, in the total number of registered patents in that province, divided by the share of registered patents in that sector at the national level, in the total number of patents in the country. If $RGTPS^{b}_{s,p} > 1$, then province p is specialized in sector s because its share in the total number of patents in the province is greater than its corresponding share at the national level. Alternatively, if $RGTPS^{b}_{s,p} < 1$, the province does not exhibit specialization.
- *Relative Geographical and Technological Patent Diversification, 'RGTPD'*. This indicator informs about the degree of sectoral diversification of a province compared with the national average. It is computed as the inverse of the sum of the absolute value of the difference between each $s = 1, \ldots, S$ sector's share in each province's

[17] Although in Model 1 (0; >0) the inclusion of '*Degree*' and '*Component Size*' is unnecessary because all observations, being isolated, are independent of each other and preferential attachment is not at work; this is not the case in subsequent models where the likelihood of developing a connection may depend on the initial degree and component size. Nevertheless, for Model 2 we decided against including these variables in the regressions because they are not useful to test the hypothesis of preferential attachment due to the lack of variability.

Table 2
Summary of regression variables, expected sign, and rationale.

Variable	Sign	Rationale
Duration (ADPP)	Undetermined (?)	Effect of monopoly power and dynamic allocative efficiency
Assignment	Undetermined (?)	Effect of owning patents with market value
Prolific	Positive (+)	Effect of having a portfolio of patents
RTD	Positive (+)	Effect of having technologically diversified patents
Firm	Negative (−)	Effect of reducing transaction costs, hidden networks, …
Importer	Negative (−)	Effect of rent-seeking behavior to innovate through technology transfer
Female	Negative (−)	Effect of gender
Non-Resident	Positive (+)	Effect of exposure to foreign innovation activity
Degree	Positive (+)	Effect of preferential attachment (degree)
Component Size	Positive (+)	Effect of belonging to large components
RGTPS	Undetermined (?)	Effect of geographical specialization
RGTPD	Undetermined (?)	Effect of geographical diversification
Patents PC	Positive (+)	Effect of innovation activity at the geographical level

Table 3
Descriptive statistics of regression variables (patent/patentee data).

Variable	Base Period (1878–1914)							Cumulative Period (1878–1939)						
	N	%	Min.	Mean	Median	Max.	St. Dev.	N	%	Min.	Mean	Median	Max.	St. Dev.
ADPP	28,015	–	0	2.35	1.89	20	2.10	73,823	–	0	2.80	2	20	2.69
Prolific	28,015	–	1	4.49	2	65	6.32	73,823	–	1	8.59	3	202	21.03
1	9332	33.31 %	1	1.00	1	1	0.00	21,604	29.26 %	1	1.00	1	1	0.00
2–5	12,476	44.53 %	2	3.02	3	5	1.07	30,049	40.70 %	2	3.05	3	5	1.07
6–10	3289	11.74 %	6	7.57	7	10	1.35	10,522	14.25 %	6	7.65	7	10	1.38
11–50	2918	10.42 %	11	19.09	17	45	8.26	9251	12.53 %	11	19.88	17	19	8.53
51–100	65	0.23 %	65	65.00	65	65	0.00	1365	1.85 %	51	71.44	67	99	15.49
>100	–	–	–	–	–	–	–	1032	1.40 %	103	153.96	149	202	31.57
RTD	28,015	–	1	2.11	1.00	15	1.69	73,823	–	1	2.27	2	15	1.91
Importer	28,015	–	0	0.18	0	1	0.31	73,823	–	0	0.19	0	1	0.32
Degree	28,015	–	0	0.34	0	13	0.88	73,823	–	0	0.48	0	36	1.64
Component Size	28,015	–	1	1.45	1	17	1.39	73,823	–	1	1.72	1	43	2.82
RGTPS	28,015	–	0.36	1.37	1.02	24.83	1.37	73,823	–	0.08	1.34	0.99	115.16	1.77
RGTPD	28,015	–	0.20	0.22	0.21	0.26	0.01	73,823	–	0.45	0.07	0.07	0.12	0.01
PatentsPC	28,015	–	1.30	66.02	76.16	107.84	42.22	73,823	–	1.56	70.08	76.16	107.84	41.45

Notes: All variables entering the regressions correspond to the base period (1878–1914).
ADPP (average duration of patents per patentee); *RTD* (relative technological diversification of the patentee); *Prolific* (patents per patentee); *Importer* (% of patents of introduction over all the patents owned by the patentee); *Degree* (number of links of the patentee); *Component Size* (size of the component to which the patentee belongs); *RGTPS* (relative geographical and technological patent specialization); *RGTPS* (relative geographical and technological patent diversification); *PatentsPC* (patents per capita).
Source: Authors' calculations.

registered patents, and its share at the national level (Duranton and Puga, 2000).[18]

$$RGTPD_p^b = 1 \Big/ \sum_{s=1}^{S} \left| Share\,RP_{s,p}^b - Share\,RP_{s,country}^b \right| \tag{4}$$

– Last, *Relative share of patents* per capita, '*PatentsPC*'. This indicator controls for the population at the provincial level and serves as a proxy of the relative market size where innovations take place. It is computed as the patents per capita in province *p* where the patentee operates, divided by the patents per capita at the national level—or, alternatively, the provincial share of patents divided by the provincial share of the population

$$PatentsPC_{s,p}^b = \frac{\sum_{s=1}^{S} RP_{s,p}^b \Big/ Pop_p^b}{\sum_{s=1}^{S}\sum_{p=1}^{P} RP_{s,p}^b \Big/ Pop_{p,country}^b} = \frac{Share\,RP_{p,country}^b}{Share\,Pop_{p,country}^b}, \quad s = 1, \dots, S,\ p = 1, \dots, P. \tag{5}$$

2) *Sector specification:* Differences among technological sectors when considering collaboration through patents are also relevant. Therefore, we control the different *Sectors* to which patents belong by defining the corresponding dummy variables, denoted by η_s in Eq. (6) below, for the different industries (see Table 1). The inclusion of these dummies intends to capture additional information related to sectoral characteristics beyond those already considered in the definition of the geographical variables.

We summarize in Table 2 the information about the explanatory variables along with the expected relationship (sign) with the likelihood of transitioning and its rationale.

Considering the compound matrix of patent-patentee interactions, Table 3 shows the descriptive statistics for the regression variables (excluding dummies). The statistics on the left-hand side correspond to

[18] '*Relative Technological Diversification, RTD*' and '*Relative Geographical and Technological Patent Diversification, RGTPD*' are also interacted in the regression with crossed effects to capture complementary interactions in the product and geographic spaces, see Tables A.2.2 and A.2.3 in Appendix 2.

the base period, thereby entering the regressions, while those reported on the right-hand side pertain to the cumulative period. The longest average duration of patents per patentee, *ADPP*, is the legal maximum of 20 years, while the mean duration increases from 2.35 years to 2.80 years. Regarding how *Prolific* the actors of the network are, most of the patentees own 5 or less patents (77.84 % and 69.96 %), while the most prolific patentees hold a total of 65 patents in the base period and 202 patents in the cumulative period. The most diverse patentee registers patents in all 15 distinct sectors (*RTD*), and the mean number of sectors is rather low—although it slightly increases from 2.11 to 2.27, which along with the scarce number of highly connected *Firms* and *Non-residents*, bears witness to the technological and institutional deficits that affected the growth and economic development of Spain. In general, the descriptive statistics are similar in the base and cumulative periods.

4.3. Econometric specification and estimation

As previously mentioned, we define different dependent binary variables related to the evolution of the degree according to models 1 thru 3. The general econometric specification is the following:

$$\begin{aligned} Y^{b,c}_{Model\#} &= \beta_0 + \beta_1 ADPP^b + \beta_2 Assignment^b + \beta_3 Prolific^b + \beta_4 RTD^b \\ &+ \beta_5 Firm^b + \beta_6 Importer^b + \beta_7 Female^b + \beta_8 Non-Resident^b \\ &+ \beta_9 Degree^b + \beta_{10} Component\,Size^b + \beta_{11} RGTPS^b_{s,p} + \beta_{12} RGTPD^b_{s,p} \\ &+ \beta_{13} PatentsPC^b_p + \eta^b_s + \varepsilon_{b,c}, \end{aligned} \tag{6}$$

where η^b_s is a dummy variable capturing sector-specific effects.

Regarding the estimation method, SNA literature offers specific models to estimate network growth controlling for the topological characteristics of the network such as the temporal exponential random graph model (TERGM) and the stochastic actor-oriented model (SAOM), whose common precursor is the multinomial logit model. However, their methodological assumptions are rather limiting, while the information requirements on the network structure are too demanding considering our available historical connectivity patterns (Leifeld and Cranmer, 2019). Thus, we decided on the linear probability model (LPM) as the one better suited among all the binary models, given our data restrictions.[19]

5. Results and discussion: Estimated effects of the transitioning models

The discussion of the estimation results is organized following three subsections, one for each of the transition models studying specific growth patterns of the innovation network, i.e., according to the specified dependent variable $Y^{b,c}_{Model\#}$, # = 1,2,3, as discussed in Section 4.1. Due to its relevance in explaining the emergence of collaboration, we discuss first the results of the '*Essential change*' Model 1: (0; >0). Then we comment on the '*Dyads and Triads*' Model 2: (1; 2, 3) ∪ (2; 3), studying the initial stages of connectivity growth. Finally, we turn to the '*High connectivity*' Model 3: (1, 2, 3; >3).

We present the estimation results for all models in Table 4.[20] Columns identified as [1.1], [2.1], and [3.1] show the results for each model in its '*Baseline specification*' including the whole sample of observations except '*Non-resident*' in [1.1].[21] For instance, column [1.1] reports the results for the first model (0; >0) representing the transition from being isolated to being connected, considering the '*Baseline specification*'. Columns [1.2], [2.2], and [3.2] correspond to the '*Geographic specification*' of each model. Finally, columns [1.3], [2.3], and [3.3] present the estimations for the '*Sector specification*'. This hierarchical presentation by model and specification eases the comparison between the different results. Also, as previously discussed, Model 3 includes the '*Degree*' and '*Component Size*' variables intended to capture the likely existence of preferential attachment and the effect of belonging to large groups.

The values of the coefficient estimates are presented for the sake of completeness. However, they cannot be directly used to interpret the magnitude of the effects of the variables on the probability of transition. To determine these magnitudes, we compare the value of the coefficients with the unconditional probability of transition in each model, i.e., the number of observed transitions divided by the number of cases that may potentially transition. For example, in the '*Baseline specification*' of Model 1 [1.1], an increment in the average duration of patents per patentee, *ADPP*, equal to its standard deviation in the base period (1878–1914), decreases the probability of transitioning by 9.07 % of the unconditional value: −9.07 % (= [−0.000844 × 2.0978/0.01955] × 100). Table 5 shows the results for these marginal effects. In what follows, when comparing the effects of the different regressors on the growth of the collaborating innovation network, we focus on these results as they are the ones that can be compared in a meaningful way.

5.1. Overcoming isolation: 'Essential change'

The first columns of Table 4 and Table 5 report the results for the models related to transitions overcoming isolation (Model 1 (0; >0)). Regarding the '*Baseline specification*' shown in column [1.1], the two first variables are, as in the rest of the models, either negative (*ADPP*) or non-significant (*Assignment*). As previously shown, the measure of duration, *ADPP*, reduces the probability of transitioning by about nine percentage points, −9.07 %. Therefore, the longer the lifespan of the patent, the lower the probability of seeking collaborations. In contrast, being a highly *Prolific* patentee encourages collaboration in a greater magnitude (11.87 %). The indicator of relative technological diversification (*RTD*) is significant in all the cases and has the expected positive value, largely fostering collaboration (57.61 %) and presenting the highest positive value of all variables in Model 1. The effect of *Firm* is significant and negative, as expected because firms are networks in themselves and, at that time, companies were not inclined to patent with individuals or other firms. In contraposition to *RTD*, *Firm* is the largest negative factor hampering collaboration in Model 1. Also, as expected, the fact that the patentee free-rides on foreign technology through 'patents of introduction' (*Importer*) does not help to overcome isolation. Nevertheless, this effect is not statistically significant. Another factor analyzed is the gender of the patentee (*Female*), which shows a negative albeit reduced effect.

Focusing now on the '*Geographical specification*' [1.2], patent provincial specialization $RGTPS^b_{p,s}$ is significant with a relevant effect of 12.4 %. This shows that geographical clustering is key to explaining the emergence of first connections among agents.[22] The positive effect of locating in a province specialized in the sector where the patentee operates is coupled, at this time of first industrialization, with the existence of a diversified environment where innovation activities take place. In this transitioning model, $RGTPD^b_{s,p}$ also presents a positive and

[19] In this respect, the main drawbacks of the LPM are addressed in this study. We estimate the model with robust standard errors (Wooldridge, 2010, chap. 4) to control for heteroskedasticity. We also check if predicted values fall outside the unit interval (probabilities larger than one or negative). Online Appendix 2 shows the reliability of the LPM model based on goodness-of-fit tests and post-estimation results.

[20] All estimations are carried out using Stata (version 15.1), https://www.stata.com/.

[21] Patents from isolated non-residents are excluded from the dataset.

[22] The geographical issue is critical in Spain and should be studied in detail in further work. Specifically, in the base period from 1878 to 1914, approximately 44 % of the patents were signed in Barcelona and 20 % in Madrid, the two main hubs of knowledge in the country.

Table 4
Regression results of the transitioning models by specifications.

Specification:	Model 1 (0;>0)			Model 2 (1;2,3) ∪ (2;3)			Model 3 (1,2,3;>3)		
	Basic	+Geo	+Sector	Basic	+Geo	+Sector	Basic	+Geo	+Sector
Variable	**[1.1]**	**[1.2]**	**[1.3]**	**[2.1]**	**[2.2]**	**[2.3]**	**[3.1]**	**[3.2]**	**[3.3]**
ADPP	−0.000844**	−0.000935**	−0.000997**	−0.00400***	−0.00354***	−0.00354***	−0.00202***	−0.00180***	−0.00189***
	(0.000406)	(0.000418)	(0.000409)	(0.00118)	(0.00123)	(0.00126)	(0.000466)	(0.000436)	(0.000481)
Assignment	0.000407	−0.000557	−0.000396	0.0283***	0.0287***	0.0280**	0.0200*	0.0203*	0.0220**
	(0.0058)	(0.00587)	(0.00585)	(0.0106)	(0.0108)	(0.011)	(0.0103)	(0.0104)	(0.0104)
Prolific	0.000487*	0.000415*	0.000411	0.00601***	0.00595***	0.00607***	0.000610***	0.000669***	0.000697***
	(0.000272)	(0.000247)	(0.000268)	(0.000808)	(0.000828)	(0.000865)	(0.000223)	(0.000247)	(0.000254)
RTD	0.0141***	0.0143***	0.0143***	0.000885	0.00154	0.000796	0.00220**	0.00244**	0.00240**
	(0.00144)	(0.00145)	(0.00149)	(0.00289)	(0.00301)	(0.00318)	(0.000953)	(0.00104)	(0.00106)
Firm	−0.0137***	−0.0145***	−0.0142***	−0.0509***	−0.0538***	−0.0475***	−0.00226*	−0.000553	−0.00105
	(0.002)	(0.00202)	(0.00207)	(0.00652)	(0.00712)	(0.00774)	(0.00134)	(0.00156)	(0.00175)
Importer	−0.00133	−0.00223	−0.00222	0.0189**	0.0166**	0.0169**	−0.000906	0.00135	0.00130
	(0.00208)	(0.00223)	(0.00223)	(0.00752)	(0.00815)	(0.00817)	(0.00207)	(0.00168)	(0.00188)
Female	−0.00613**	−0.00647**	−0.00606*	−0.0135***	−0.0130***	−0.0107**	0.02400	0.00177	0.00214
	(0.00305)	(0.00324)	(0.00334)	(0.00326)	(0.00365)	(0.00468)	(0.0159)	(0.00157)	(0.00192)
Non-Resident	#	#	#	−0.0143***	–	–	−0.00588**	–	–
	#	#	#	(0.00446)	–	–	(0.00271)	–	–
Degree	–	–	–	–	–	–	0.0290***	0.0299***	0.0297***
	–	–	–	–	–	–	(0.00457)	(0.00478)	(0.00481)
Component Size	–	–	–	–	–	–	−0.00123**	−0.00122**	−0.00122**
	–	–	–	–	–	–	(0.000517)	(0.000546)	(0.000535)
$RGTPS^{b}_{s,p}$	–	0.00181***	0.00181**	–	−0.00287*	−0.00604**	–	−0.000241	0.000134
	–	(0.000684)	(0.000805)	–	(0.00169)	(0.003)	–	(0.000473)	(0.000674)
$RGTPD^{b}_{s,p}$	–	0.180**	0.153**	–	0.543***	0.669***	–	0.430***	0.423***
	–	(0.0708)	(0.0725)	–	(0.163)	(0.172)	–	(0.101)	(0.103)
$PatentsPC^{b}_{p}$	–	6.83E-05***	6.39E-05***	–	−0.0000506	−0.0000477	–	0.0000104	0.0000139
	–	(0.0000235)	(0.0000239)	–	(0.0000641)	(0.0000652)	–	(0.0000298)	(0.0000298)
Aeronautics	–	–	BASE	–	–	BASE	–	–	BASE
Agriculture / Farming	–	–	0.00258	–	–	0.0606***	–	–	0.0255**
			(0.0197)			(0.0234)			(0.0115)
Arms Industry	–	–	0.00675	–	–	0.0767**	–	–	0.00653
			(0.0205)			(0.0388)			(0.00513)
Basic Metals	–	–	0.0154	–	–	0.0440**	–	–	0.0322**
			(0.0202)			(0.0206)			(0.0127)
Chemical	–	–	0.00905	–	–	0.00326	–	–	0.0169***
			(0.0195)			(0.00789)			(0.00607)
Communications	–	–	−0.0127	–	–	0.0485	–	–	−0.00350
			(0.0218)			(0.0300)			(0.00538)
Construction	–	–	0.00451	–	–	0.0569***	–	–	0.00498
			(0.0195)			(0.0146)			(0.00303)
Electricity	–	–	0.0104	–	–	0.0302*	–	–	0.00866
			(0.0203)			(0.0155)			(0.00625)
Gas/Lighting	–	–	−0.00118	–	–	0.00172	–	–	0.00465
			(0.0201)			(0.00981)			(0.00314)
Lumber	–	–	0.00456	–	–	0.0292	–	–	0.00332
			(0.0211)			(0.0203)			(0.00391)
Machinery / Equipment	–	–	0.00244	–	–	0.0214**	–	–	0.0155***
			(0.0192)			(0.00835)			(0.00432)
Mining / Coal	–	–	0.0077	–	–	0.00280	–	–	0.00949***
			(0.022)			(0.0164)			(0.0036)
Non-rail transportation	–	–	0.0188	–	–	−0.00151	–	–	0.0208**
			(0.0208)			(0.0100)			(0.0103)
Paper / Graphic Arts	–	–	0.00012	–	–	0.00697	–	–	0.0179**
			(0.0194)			(0.0107)			(0.00763)
Railway	–	–	0.00197	–	–	0.0355**	–	–	0.0253**
			(0.0203)			(0.0156)			(0.0127)
Sea Transportation	–	–	0.0134	–	–	0.0124	–	–	0.0143
			(0.0221)			(0.0130)			(0.0116)
Services	–	–	0.00908	–	–	0.0362***	–	–	0.0167***
			(0.0192)			(0.00831)			(0.0044)
Textile	–	–	0.00492	–	–	0.00641	–	–	0.0115***
			(0.0191)			(0.0098)			(0.00332)
Food, Beverages & Tob.	–	–	0.0042	–	–	0.0219**	–	–	0.0114***
			(0.0193)			(0.00952)			(0.00438)
Constant	−0.01084***	−0.0565***	−0.0565**	−0.0142***	0.113***	0.121***	−0.0307***	−0.128***	−0.141***
Observations	21,637	20,809	20,747	5862	5526	5519	6094	5752	5744
Positive Cases	423	423	420	205	205	205	54	54	54
R-squared	0.036	0.037	0.037	0.09	0.092	0.099	0.039	0.043	0.047

Notes: Robust standard errors in parenthesis; *, **, and *** represent significance at 10 %, 5 %, and 1 % respectively.
Source: Authors' calculations.

Table 5
Estimated effects of the transitioning models by specifications.

	Model 1 (0;>0)			Model 2 (1;2,3) U (2;3)			Model 3 (1,2,3;>3)		
Specification:	**Basic**	**+Geo**	**+Sector**	**Basic**	**+Geo**	**+Sector**	**Basic**	**+Geo**	**+Sector**
Variable	**[1.1]**	**[1.2]**	**[1.3]**	**[2.1]**	**[2.2]**	**[2.3]**	**[3.1]**	**[3.2]**	**[3.3]**
ADPP	-9.07%	-9.82%	-11.00%	-24.02%	-20.04%	-20.01%	-47.87%	-40.26%	-42.22%
Assignment	~	~	~	14.03%	13.41%	13.07%	39.14%	37.49%	40.58%
Prolific	11.87%	9.78%	~	34.54%	32.24%	32.85%	13.84%	14.32%	14.90%
RTD	57.61%	57.38%	57.11%	~	~	~	9.50%	9.95%	9.77%
Firm	-17.87%	-18.15%	-18.07%	-16.35%	-16.30%	-14.37%	-2.87%	~	~
Importer	~	~	~	16.75%	13.87%	14.10%	~	~	~
Female	-3.87%	-4.00%	-3.90%	-4.77%	-4.33%	-3.56%	~	~	~
Non-Resident	#	#	#	-1.74%	-	-	-2.83%	-	-
Degree	-	-	-	-	-	-	28.96%	28.18%	27.95%
Component Size	-	-	-	-	-	-	-3.87%	-3.62%	-3.61%
$RGTPS^{b}_{z,p}$	-	12.40%	12.37%	-	-10.60%	-22.28%	-	~	~
$RGTPD^{b}_{z,p}$	-	13.06%	11.06%	-	21.22%	26.12%	-	66.41%	65.24%
$PatentsPC^{b}_{p}$	-	14.42%	13.45%	-	~	~	-	~	~
Aeronautics	-	-	BASE	-	-	BASE	-	-	BASE
Agriculture / Farming	-	-	~	-	-	26.17%	-	-	43.51%
Arms industry	-	-	~	-	-	30.52%	-	-	~
Basic Metals	-	-	~	-	-	20.50%	-	-	59.29%
Chemical	-	-	~	-	-	~	-	-	41.81%
Communications	-	-	~	-	-	~	-	-	~
Construction	-	-	~	-	-	32.38%	-	-	~
Electricity	-	-	~	-	-	14.79%	-	-	~
Gas / Lighting	-	-	~	-	-	~	-	-	~
Lumber	-	-	~	-	-	~	-	-	~
Machinery / Equipment	-	-	~	-	-	21.37%	-	-	61.17%
Mining / Coal	-	-	~	-	-	~	-	-	9.81%
Non-Rail Transportation	-	-	~	-	-	~	-	-	35.42%
Paper / Graphic Arts	-	-	~	-	-	~	-	-	40.52%
Railway	-	-	~	-	-	13.70%	-	-	38.56%
Sea Transportation	-	-	~	-	-	~	-	-	~
Services	-	-	~	-	-	38.95%	-	-	71.00%
Textile	-	-	~	-	-	~	-	-	43.65%
Food, Beverages & Tob.	-	-	~	-	-	17.96%	-	-	36.95%
Observations	21,637	20,809	20,747	5,862	5,526	5,519	6,094	5,752	5,744
Positive Cases	423	423	420	205	205	205	54	54	54
Uncond. Prob. of Trans.	1.95%	2.03%	2.02%	3.50%	3.71%	3.71%	0.89%	0.94%	0.94%

Notes: Marginal effects measure the percentage variation in the probability of transition to the unconditional probability of transition (last row). Figures not reported (∼) are not statistically significant as indicated in Table 4. The length of the color bars represents the relative (proportional) value of each variable compared to others within each specification.
Source: Authors' calculations

significant impact by increasing collaboration by up to 13.06 %, thereby in the same order of magnitude as its specialization counterpart. All these positive effects are reinforced by the market size. The estimated effect of $PatentsPC^{b}_{p}$ shows that the higher the number of patents per capita in relative terms to the national average, the larger the likelihood of increasing connections, which is similar in magnitude to the previous two geographical variables, 14.42 %.

Finally, and quite surprisingly, the dummies in the '*Sectoral specification*' [1.3] are not significant. A possible explanation is that all relevant information for this first '*Essential change*' model is already embedded in the geographical variables, which include the sectoral distribution of patents across locations. Nevertheless, this result suggests that the collaboration dynamics to overcome isolation were independent of specific technological fields, reinforcing the idea that the main incentives to establish links lay in the general necessity of financial or technical support.

5.2. '*Dyads and Triads*' model

Columns [2.1] to [2.3] in Table 4 and Table 5 show the coefficient estimations and the marginal effects related to the growth of '*Dyads and Triads*' in the innovation network (Model 2; (1; 2, 3) ∪ (2; 3)). As in the previous model, for the '*Baseline specification*', we observe first that the magnitude of duration, *ADPP*, remains negative doubling its effect for the '*Dyads* and *Triads*' model: −20.01 % in column [2.3]. Now this variable is the main negative covariate of the model. However, a similar effect is observed in how *Prolific* the patentee is, whose positive effect triples from the initial '*Essential change*' model. Specifically, the value of this variable increases from 11,87 % in [1.1] to 34.54 % in [2.1]. In contrast, a substantial reduction is observed in the coefficients of relative technological diversification (*RTD*), whose values in Table 4 are two orders of magnitude smaller than in Model 1 and lose statistical significance. Additionally, *Firms* keep hampering growth in collaboration, maintaining its negative effect at similar levels, −16.35 % in [2.1]. Regarding 'patents of introduction'—and contrary to our expectations—*Importer* now emerges as a catalyzer of collaboration: 16.75 % in [2.1]. *Female* remains about the same value at −4.77 % in [2.1], while *Non-Resident* (at −1.74 %) does not increase the likelihood of increasing connectivity—supporting the notion that the Spanish network lacked openness to foreign ideas (see the last stylized fact above).

Considering the '*Geographic specification*', we observe a reversal in the sign of the geographic and technological specialization $RGTPS^{b}_{p,s}$ from positive to negative, while $PatentsPC^{b}_{p}$ turns non-significant to explain transitions to higher connectivity. This suggests that once the basic connectivity threshold represented by Model 1 (0; >0) is met, when considering the step increments in connectivity in Model 2 the overall positive effects of the three locational variables wane. Only the positive effect of geographic and technological diversification remains. Finally, the inclusion of sectoral dummies [2.3] shows that the probability of transition to these relatively small components increases in several sectors, with remarkable values in services (38.95 %), followed by construction (32.38 %) and the arms industry (30.52 %), where collaboration to innovate could be key. However, the probability of transition also grows for other sectors technically less complex although economically very relevant during the first industrialization processes, especially in latecomers, such as '*Agriculture/Farming*' (26.17 %), as well as its processed byproducts in the sector of '*Food, Beverages & Tobacco*' (17.96 %).

5.3. 'High connectivity' model

The results corresponding to the three specifications of the model studying transitions to all possible higher degrees, Model 3 (1, 2, 3; >3), are reported in columns [3.1], [3.2], and [3.3] of Table 4 (coefficient estimates) and Table 5 (estimated effects). Determining the factors that contribute to the development of highly connected individuals is central to the emergence of densely connected networks. As before, the larger the lifespan of the patents, *ADPP*, the lower the incentives to collaborate. The increase in the negative value of the variable (−47.87 % in column [3.1]) compared to those of the previous models, greatly reinforces the result that the monopoly power granted by patents hampers collaboration. Here, the variable *Assignment* turns significant and positive (as expected, being an indication of the commercial value of patents), while being a *Firm*, *Importer*, or *Female* loses significance (albeit firms continue to be detrimental to external collaboration growth as innovation usually takes place in-house). Also, *Non-Resident* turns significant, reaching a negative value of −2.83 %.

Although preferential attachment seems to foster connectivity in the entire network, we use this model to reliably test the presence of this feature by considering all potential transitions to degrees >3. The positive and significant value of *Degree*—approximately 30 % in all three models and second only to *Assignment*—confirms that the higher the degree in the base period, the larger the probability of increasing connections, supporting the hypothesis that preferential attachment is at work. However, the variable *Component Size* presents a negative and significant value (−3.87 %). Thus, belonging to larger components in the base period does not ensure increased collaboration levels and higher degrees. These results reinforce the notion—exposed in Section 4.2 when commenting on this variable—that there exist counteracting dynamics against connection increases. Collaboration has both benefits and costs and, in this case, the technological and economic opportunities brought by the membership to larger components did not compensate for the efforts of establishing and maintaining new links. Hence, we hypothesize that the Spanish institutional and historical scenario negatively affected the formation of larger and more solid hubs. This must be confirmed by further studies on the effect of membership size in other innovation networks.

The inclusion of the geographical variables qualifies previous results obtained for the '*Baseline specification*' for the variable of relative diversification ($RGTPD^{b}_{s,p}$), whose value increases to 66.41 %, while the other two ($RGTPS^{b}_{p,s}$ and $PatentsPC^{b}_{p}$) are, once again, non-significant. Finally, the '*Sectoral specification*' confirms the robustness of the sectoral dummies mentioned in the previous model, especially in low-tech activities such as '*Agriculture/Farming*', '*Food, Beverages & Tobacco*', or '*Services*', but also key heavy industries such as '*Basic Metals*', '*Machinery/Equipment*'; or '*Railway*', that strongly grew in Spain after 1914.

5.4. Robustness checks

We conclude this section by referring the reader to online Appendix 2 where we provide a general evaluation of the goodness-of-fit and post-estimation results of the above regressions. Besides this, we have also run two additional models to determine the robustness of the results. First, we consider a '*High connectivity*' model similar to Model 3 but considering transitions from 1 or 2 degrees to >2; i.e., Model 3′ (1, 2; >2). Second, we also check if the previous results for all three complete specifications (+Sector) in Models 1, 2, and 3 are robust to the introduction of crossed-effects among the basic variables to account for possible interactions among them. We find that results for these two additional specifications confirm the sign, value, and significance of the variables in the three models discussed in the previous sections. A discussion of the estimated results, including regression coefficients and marginal effects, can be found in online Appendix 2 (see Tables A.2.2. and A.2.3).

6. Conclusions

Collaboration is a key issue in fostering knowledge creation and innovation processes. During the last decades, studies on co-authors of scientific articles or co-inventors of new technologies have grown exponentially, reflecting the increasing interest of scholars in the causes of cooperation and their effects. Several of those researchers rely on SNA methods and patent data to explore the structural dynamics of innovation networks. Their findings demonstrate that increasing network connectivity reinforces access to information, creativity, technical cross-fertilization, and collaborative solutions to innovate, while isolation works oppositely. In general, this literature focuses on contemporary cases in the most advanced economies and, therefore, on mature networks.

In a novel way, this work combines historical patent files, SNA techniques, and discrete choice regressions to delve into the origins and evolution of collaboration and innovation networks in Spain, a country with traditional backwardness in R&D production. Thus, we focus the investigation on the emerging state of a network, when isolation and disconnection prevail. The main goal is to study the factors that initially triggered or hampered collaboration in the past because those first steps are critical to understanding the long-term divergence in cooperation structures and performance among distinct national innovation systems. Hence, the analysis provides new and fresh insights regarding the dynamics of innovation in less developed economies, while providing relevant policy implications for both technologically dependent areas and innovative leading countries.

The research methodology begins with the construction of the historical innovation network and its detailed topological analysis, based on patents registered in Spain from 1878 to 1939, during a key stage of the country's modernization process. During this period, Spain had a hybrid patent regime that met the basics of the 1883 Paris Convention for the international protection of IPRs but maintained early-nineteenth-century characteristics to foster industrialization and technology transfer such as patents of introduction or compulsory working clauses. Then, we study the transitions in the actors' links (degree) from the first period (1878–1914) into the cumulative period (1878–1939) by specifying and estimating a linear probability model on the explanatory factors. We analyze three models (with three distinct specifications each, progressively introducing geographical and sectoral variables): 1) '*Essential change*', from being isolated to establishing links; 2) '*Dyads and Triads*', the evolution from 1 to 2 connections to 2–3 links; and 3) '*High connectivity*', the transition of connected actors to 4 or more links.

The conclusions are relevant and straightforward. The literature on SNA and innovation highlights the key role of increasing network density and openness—i.e. external collaboration—to innovation. However, the topological analysis of the Spanish case shows a scattered and disconnected network with a modest evolution without very large or giant components. Although this may be common to the first steps of innovation networks in the past, the results for Spain confirm a critically slow growth of connectivity throughout time compared to other successful innovative countries (see Andersson et al., 2019; van der Wouden, 2020) and, especially, a closeness to foreign collaboration in R&D through common patents (just 116 patents linked residents with non-residents). The absence of foreign interactions prevents not only the quick penetration of new scientific and technical information but also the cross-fertilization and merging with local ideas.[23] Therefore, the lack of collaboration density and openness seems a key issue for the understanding of the backwardness of an innovation system. Indeed, recent studies focusing on current technologically peripheral countries—particularly within the Latin American region—have unveiled similar characteristics to the Spanish case a century ago. Most of these developing countries show sparse patent networks, with highly fragmented structures and no giant components (Bianchi et al., 2023). However, unlike Spain in the past, today's Latin American networks evolved towards a clear outward orientation, increasing their connections with inventors in other regions. This could be explained by the current globalization process—which facilitates cross-border interactions and knowledge exchange—and, above all, by the methodology used to build these patent networks. The study extracts Latin American patentees/inventors from U.S. patents, i.e., from the first worldwide market for technology where international collaboration is more plausible. Although further research on backward regions is needed, the results confirm that weak innovation systems are characterized by less densified R&D/patent networks and lack of cooperation. Thus, understanding the mechanism that drive collaboration to innovate is critical.

In our investigation, the analysis of degree transitions within the Spanish network—although it is slightly nuanced according to each model and specification—provides interesting findings concerning the explanatory factors to cooperate: First, there is a consistent negative effect of patents' duration on collaboration (*ADPP*), outweighing the possible positive effects attributed to the quality of the inventions they protect. This is a relevant issue because significant economic literature has reflected on distinct policies concerning optimal patent length and breath and their consequences on R&D management (see, e.g. Gilbert and Shapiro, 1990; Scotchmer and Green, 1990; O'Donoghue et al., 1998) and several resolutely recommended to shorten patent length, especially if market size increases in time (Boldrin and Levine, 2008, 2009, 2013). Our work reveals a new and unintended consequence of intellectual monopolies: they hamper collaboration to innovate. Indeed, scholars have also highlighted that increasing market power (as that granted by patents) diminishes incentives to innovate individually or through joint ventures (Cabral, 2000, chap. 16; Motta, 2004, chap. 2). Therefore, assuming that collaboration is a prerequisite to achieve successful innovation systems over the long term, then shorter patents and lower barriers to knowledge use and recombination may facilitate R&D cooperation and increase innovation results. Further research is undoubtedly needed to confirm this.

Second, being a firm brings a consistently negative effect on collaboration in most models and specifications. A company is a kind of 'hidden network' itself and, at least in the past, they seem to reject external collaboration with third parties, as innovation activities taking place inside the firm reduce transaction costs. Furthermore, evidence from US R&D companies reveals that the propensity to directly patent and employ researchers in their R&D departments—as a way to internalize collaboration—dramatically increased during the first third of the twentieth century (Nicholas, 2011b). This signals the path towards a 'corporate' patent system where cooperation mainly occurs within the firm's R&D sections, although—nowadays—it can also spread either through research ventures among corporations or through interactions of (free-lance) scientist/inventors working for distinct companies, which gives place to the mature patent networks that the literature analyzes.

Third, and in contrast, another set of variables increases the probability of transitioning to higher degrees and favors collaboration. That is the case of the patentee's productivity, measured by the number of patents (*Prolific*), positive in all the models and specifications. Likewise, the patentee's technological diversification, measured by the distinct sectors where they patent (*RTD*) has a positive effect (although it loses significance in Model 2). Hence, more patent applications in more diversified fields benefit collaboration (corroborated by the crossed effect of these two variables in Tables A.2.2 and A.2.3 in online Appendix 2). In essence, the positive sign of innovation productivity and, especially, *RTD* suggests that higher levels of sectoral diversification correspond to an enhanced ability to engage with diverse knowledge, i.e. an increasing *absorptive capacity* that favors cooperation and more complex innovations. 'Patents of introduction' (*Importer*)—a form of protection that essentially allows free-riding on foreign technologies—results in small increments of degree just in Model 2, reflecting an ambiguous effect more related to technology copying than to invention activities. This kind of patent was common in almost all countries during the nineteenth century but was progressively abandoned to reinforce the protection of true inventors. Spain maintained patents of introduction and other weak characteristics (lack of previous technical exams, etc.) until 1986, with the intention of promoting innovation through technology transfer. In light of the true nature of the increased cooperation—aimed at copying rather than inventing—this institutional setting seems to have weakened the R&D system by deepening the country's scientific and technological dependence over the long term. Additionally, these kinds of institutions may generate socio-cultural dynamics of collaboration that do not promote creative effort, Schumpeterian entrepreneurship, or scientific meritocracy, which characterize developed R&D systems. These attitudes are historically path-dependent and difficult to reverse. For instance, in the period under analysis there were few institutions, associations or specialized publications in Spain aimed at promoting science, inventive activity, and research (and those that did exist had little social or political relevance). However, these kinds of organizations and publications were key and socially prominent hubs for connecting researchers, professionals, and companies in pioneering countries, as was the case with the British *Royal Society*, the French *Académie des Sciences* or the journal *Scientific American* in the United States, to cite a few examples.

Fourth, the geographical analysis shows that to overcome isolation (Model 1) and regarding the location where the patentee operates, both provincial sectoral specialization (*RGTPS*) and diversification (*RGTPD*) have positive effects on the probability of connecting. However—supporting what the literature shows for mature networks (Cantner and Graf, 2004; Breschi and Lenzi, 2015; van der Wouden and Rigby, 2019)—once connected (Models 2 and 3) spatial specialization does not influence and—in tune with our previous findings regarding patentee's technological diversification (*RTD*)—the geographical diversification affects positively (again, this is confirmed in Tables A.2.2 and A.2.3 when we analyze the crossed effect of *RTD* × *RGTPD*). Finally, the

[23] This lack of openness seems to be a long-term characteristic of the Spanish R&D system. For instance, the percentage of foreign scholars in Spanish universities in 2017 was the 2.1 % compared to 13.5 % in Sweden, 17.5 % in Denmark; 27 % in the United Kingdom, or 43 % in Switzerland (European Commission, 2017, p. 103)

existence of large markets, proxied by population (*PatentsPC*), favors cooperation at least to abandon isolation (Model 1), while operating in specific sectors has also a positive effect on being increasingly connected after overcoming isolation. That is the case of low-tech but significant economic activities at the time, such as agriculture and farming; food, beverages, and tobacco industries; or the service sector; as well as heavier key industries such as the production of basic metals; machinery and equipment; or the railway.

Finally, once demonstrated that emergent collaboration follows preferential-attachment paths (through the inclusion of a *Degree* covariate for measuring the effect of the actors' previous links in increased connectivity), we also expected a positive effect of the size of the component to which patentees belong on the probability of transitioning (as the larger the component the greater the potential flow of knowledge, information, and opportunities). However, in this case, 'size does not matter'. At least in the Spanish emergent innovation network, being a patentee in a large component does not have positive effects on connecting people. This suggests that the expected marginal benefits of increased connectivity (due to the possible gains in access to information or positioning in the network) did not compensate for the incurred costs (due to the weaknesses of the institutional environment). This duality aligns well with our results above on how the negative effect on cooperation of strong patent monopolies can nullify the advantages of being the inventor of pioneering and relevant technologies. These opposing forces are intricately tied to institutional contexts and power dynamics, encompassing both positional interest within the network (where larger components assume greater centrality) and exclusionary capacities (stemming from enduring IPRs). Although all these results are robust, it is necessary to conduct further research in other innovation networks and contexts to confirm our findings.

To sum up, if cooperation is critical to improving the scale and scope of R&D activities elsewhere, and especially in backward countries or regions, then our findings suggest, from a research policy-making perspective, to reflect on the impact of excessive duration of IPRs and also on institutional weaknesses (as effective and short protection may enhance collaboration); to encourage multidisciplinary thinking and diversity, both technologically and geographically (including international dimensions); and to carefully analyze and balance both the entrepreneurial and independent environments regarding R&D and IPRs (as corporations strongly influence the patterns of collaboration structures).

CRediT authorship contribution statement

Sergio Barbosa: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing, Project administration, Visualization. **Patricio Sáiz:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **José L. Zofío:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data is available at https://www.ibcnetwork.org/e_research_resource.php?id=3 and http://historico.oepm.es/patentes.php.

Acknowledgements

The authors express their gratitude to three anonymous referees whose comments and suggestions substantially helped to improve the quality of the study. Likewise, the authors acknowledge: i) useful comments from Prof. Jochen Streb and attendees to the 2023 Economic History Seminars of the University of Mannheim where this study was presented; and ii) key insights and methodological assistance from Pablo Galaso, Sergio Palomeque, and their research team at the *Instituto de Economia* of Universidad de Montevideo. José L. Zofío thanks the financial support from grants PID2022-136383NB-I00 and PID2022-138212NA-I00, funded by MCIN/AEI/ 10.13039/501100011033 and by the European Regional Development Fund, "A way of making Europe".

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.respol.2024.104990.

References

Ahuja, G., 2000. Collaboration networks, structural holes, and innovation: a longitudinal study. Adm. Sci. Q. 45, 425–455. https://doi.org/10.2307/2667105.
Almeida, P., Hohberger, J., Parada, P., 2011. Individual scientific collaborations and firm-level innovation. Ind. Corp. Chang. 20, 1571–1599. https://doi.org/10.1093/icc/dtr030.
Andersson, D.E., Galaso, P., Sáiz, P., 2019. Patent collaboration networks in Sweden and Spain during the second industrial revolution. Ind. Innov. 26, 1075–1102. https://doi.org/10.1080/13662716.2019.1577720.
Axelrod, R.M., 2006. The Evolution of Cooperation. Basic Books, New York.
Baba, Y., Shichijo, N., Sedita, S.R., 2009. How do collaborations with universities affect firms' innovative performance? The role of "Pasteur scientists" in the advanced materials field. Res. Policy 38, 756–764. https://doi.org/10.1016/j.respol.2009.01.006.
Balconi, M., Breschi, S., Lissoni, F., 2004. Networks of inventors and the role of academia: an exploration of Italian patent data. Res. Policy 33, 127–145. https://doi.org/10.1016/S0048-7333(03)00108-2.
Barabási, A.-L., 2003. Linked: How Everything Is Connected to Everything Else and What it Means for Business, Science, and Everyday Life. Plume, New York.
Barabási, A.-L., 2010. Bursts: The Hidden Pattern behind Everything we Do. Dutton, New York.
Barabási, A.-L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512. https://doi.org/10.1126/science.286.5439.509.
Bettencourt, L.M.A., Lobo, J., Strumsky, D., 2007. Invention in the city: increasing returns to patenting as a scaling function of metropolitan size. Res. Policy 36, 107–120. https://doi.org/10.1016/j.respol.2006.09.026.
Bianchi, C., Galaso, P., Palomeque, S., 2021a. Patent collaboration networks in Latin America: extra-regional orientation and core-periphery structure. J. Dent. Sci. Res. 10, s59–s70. https://doi.org/10.5530/jscires.10.1s.22.
Bianchi, C., Galaso, P., Palomeque, S., 2021b. The trade-offs of brokerage in inter-city innovation networks. Reg. Stud. 1–14 https://doi.org/10.1080/00343404.2021.1973664.
Bianchi, C., Galaso, P., Palomeque, S., 2023. Absorptive capacities and external openness in underdeveloped innovation systems: a patent network analysis for Latin American countries 1970–2017. Camb. J. Econ. (forthcoming). https://doi.org/10.1093/cje/bead034.
Bianconi, G., Darst, R.K., Iacovacci, J., Fortunato, S., 2014. Triadic closure as a basic generating mechanism of communities in complex networks. Phys. Rev. E 90, 042806. https://doi.org/10.1103/PhysRevE.90.042806.
Boldrin, M., Levine, D.K., 2008. Against Intellectual Monopoly. Cambridge University Press, Cambridge.
Boldrin, M., Levine, D.K., 2009. Market size and intellectual property protection. Int. Econ. Rev. 50, 855–881. https://doi.org/10.1111/j.1468-2354.2009.00551.x.
Boldrin, M., Levine, D.K., 2013. The case against patents. J. Econ. Perspect. 27, 3–22. https://doi.org/10.1257/jep.27.1.3.
Boschma, R., Balland, P.-A., Kogler, D.F., 2015. Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. Ind. Corp. Chang. 24, 223–250. https://doi.org/10.1093/icc/dtu012.
Breschi, S., Lenzi, C., 2015. The role of external linkages and gatekeepers for the renewal and expansion of US cities' knowledge base, 1990–2004. Reg. Stud. 49, 782–797. https://doi.org/10.1080/00343404.2014.954534.
Breschi, S., Lenzi, C., 2016. Co-invention networks and inventive productivity in US cities. J. Urban Econ. 92, 66–75. https://doi.org/10.1016/j.jue.2015.12.003.
Cabral, L.M.B., 2000. Introduction to Industrial Organization. MIT Press, Cambridge, Mass.
Cantner, U., Graf, H., 2004. Cooperation and specialization in German technology regions. J. Evol. Econ. 14, 543–562. https://doi.org/10.1007/s00191-004-0229-5.

Cantner, U., Graf, H., 2006. The network of innovators in Jena: an application of social network analysis. Res. Policy 35, 463–480. https://doi.org/10.1016/j.respol.2006.01.002.

Cantner, U., Meder, A., Ter Wal, A.L.J., 2010. Innovator networks and regional knowledge base. Technovation 30, 496–507. https://doi.org/10.1016/j.technovation.2010.04.002.

Cantner, U., Graf, H., Herrmann, J., Kalthaus, M., 2016. Inventor networks in renewable energies: the influence of the policy mix in Germany. Res. Policy 45, 1165–1184. https://doi.org/10.1016/j.respol.2016.03.005.

Carnabuci, G., Operti, E., 2013. Where do firms' recombinant capabilities come from? Intraorganizational networks, knowledge, and firms' ability to innovate through technological recombination. Strateg. Manag. J. 34, 1591–1613. https://doi.org/10.1002/smj.2084.

Coffano, M., Foray, D., Pezzoni, M., 2017. Does inventor centrality foster regional innovation? The case of the Swiss medical devices sector. Reg. Stud. 51, 1206–1218. https://doi.org/10.1080/00343404.2017.1329585.

de Araújo, I.F., Gonçalves, E., Taveira, J.G., 2019. The role of patent co-inventorship networks in regional inventive performance. Int. Reg. Sci. Rev. 42, 235–280. https://doi.org/10.1177/0160017618770717.

del Fresno, M., 2015. Connecting the disconnected: social work and social network analysis. A methodological approach to identifying network peer leaders. Arbor 191, a209.

Drivas, K., 2022. The role of technology and relatedness in regional trademark activity. Reg. Stud. 56, 242–255. https://doi.org/10.1080/00343404.2020.1808883.

Duranton, G., Puga, D., 2000. Diversity and specialisation in cities: why, where and when does it matter? Urban Stud. 37 https://doi.org/10.1080/0042098002104.

Endres, A.M., Harper, D.A., 2019. Economic development and complexity: the role of recombinant capital. Camb. J. Econ. 44, 157–180. https://doi.org/10.1093/cje/bez036.

Erdős, P., Rényi, A., 1959. On random graphs I. Publ. Math. Debr. 6, 290–297. https://doi.org/10.5486/PMD.1959.6.3-4.12.

European Commission, 2017. Modernisation of Higher Education in Europe: Academic Staff (Eurydice Report). Publication Office of the European Union, Luxembourg.

Fleming, L., Frenken, K., 2007. The evolution of inventor networks in the Silicon Valley and Boston regions. Adv. Complex Syst. 10, 53–71. https://doi.org/10.1142/S0219525907000921.

Fleming, L., King, C., Juda, A.I., 2007. Small worlds and regional innovation. Organ. Sci. 18, 938–954. https://doi.org/10.1287/orsc.1070.0289.

Fritsch, M., Kauffeld-Monz, M., 2010. The impact of network structure on knowledge transfer: an application of social network analysis in the context of regional innovation networks. Ann. Reg. Sci. 44, 21–38. https://doi.org/10.1007/s00168-008-0245-8.

Galaso, P., Kovářík, J., 2021. Collaboration networks, geography and innovation: local and national embeddedness. Pap. Reg. Sci. 100, 349–377. https://doi.org/10.1111/pirs.12578.

Gazni, A., Sugimoto, C.R., Didegah, F., 2012. Mapping world scientific collaboration: authors, institutions, and countries. J. Am. Soc. Inf. Sci. Technol. 63, 323–335. https://doi.org/10.1002/asi.21688.

Gilbert, R., Shapiro, C., 1990. Optimal patent length and breadth. RAND J. Econ. 21, 106–112. https://doi.org/10.2307/2555497.

Goossen, M.C., Paruchuri, S., 2022. Measurement errors and estimation biases with incomplete social networks: replication studies on intra-firm inventor network analysis. Res. Policy 51, 104404. https://doi.org/10.1016/j.respol.2021.104404.

Graf, H., 2011. Gatekeepers in regional networks of innovators. Camb. J. Econ. 35, 173–198. https://doi.org/10.1093/cje/beq001.

Graf, H., 2012. Inventor networks in emerging key technologies: information technology vs. semiconductors. J. Evol. Econ. 22, 459–480. https://doi.org/10.1007/s00191-011-0248-y.

Graf, H., Henning, T., 2009. Public research in regional networks of innovators: a comparative study of four east German regions. Reg. Stud. 43, 1349–1368. https://doi.org/10.1080/00343400802251460.

Graf, H., Krüger, J.J., 2011. The performance of gatekeepers in innovator networks. Ind. Innov. 18, 69–88. https://doi.org/10.1080/13662716.2010.528932.

Granovetter, M.S., 1973. The strength of weak ties. Am. J. Sociol. 78, 1360–1380.

Griliches, Z., 1990. Patent statistics as economic indicator. J. Econ. Lit. 28, 1661–1707.

Harhoff, D., Wagner, S., 2009. The duration of patent examination at the European patent office. Manag. Sci. 55, 1969–1984. https://doi.org/10.1287/mnsc.1090.1069.

Innocenti, N., Capone, F., Lazzeretti, L., 2020. Knowledge networks and industrial structure for regional innovation: an analysis of patents collaborations in Italy. Pap. Reg. Sci. 99, 55–72. https://doi.org/10.1111/pirs.12478.

Jackson, M.O., 2010. Social and Economic Networks. Princeton University Press, Princeton, N.J.

Jackson, M.O., Rogers, B.W., 2007. Meeting strangers and friends of friends: how random are social networks? Am. Econ. Rev. 97, 890–915. https://doi.org/10.1257/aer.97.3.890.

Jaffe, A., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Q. J. Econ. 108, 577–598.

Kali, R., 2003. The city as a giant component: a random graph approach to Zipf's law. Appl. Econ. Lett. 10, 717–720. https://doi.org/10.1080/1350485032000139006.

Katz, R., Allen, T.J., 1982. Investigating the not invented here (NIH) syndrome: a look at the performance, tenure, and communication patterns of 50 R & D project groups. RD Manag. 12, 7–20. https://doi.org/10.1111/j.1467-9310.1982.tb00478.x.

Khan, B.Z., 2013. Selling ideas: an international perspective on patenting and markets for technological innovations, 1790–1930. Bus. Hist. Rev. 87, 39–68. https://doi.org/10.1017/S0007680513000135.

Lehmann-Hasemeyer, S., Streb, J., 2020. Discrimination against foreigners: the Wuerttemberg patent law in administrative practice. J. Econ. Hist. 80, 1071–1100. https://doi.org/10.1017/S0022050720000479.

Leifeld, P., Cranmer, S.J., 2019. A theoretical and empirical comparison of the temporal exponential random graph model and the stochastic actor-oriented model. Netw. Sci. 7, 20–51. https://doi.org/10.1017/nws.2018.26.

Lerner, J., 2000. 150 years of patent protection (working paper no. 7478). National Bureau of economic research.

Lerner, J., 2005. 150 years of patent office practice. Am. Law Econ. Rev. 7, 112–143.

Lobo, J., Strumsky, D., 2008. Metropolitan patenting, inventor agglomeration and social networks: a tale of two effects. J. Urban Econ. 63, 871–884. https://doi.org/10.1016/j.jue.2007.07.005.

Mauleón, E., Bordons, M., 2014. Indicadores de actividad tecnológica por género en España a través del estudio de patentes europeas. Rev. Esp. Doc. Científica 37, e043. https://doi.org/10.3989/redc.2014.2.1093.

Menzel, M.-P., Feldman, M.P., Broekel, T., 2017. Institutional change and network evolution: explorative and exploitative tie formations of co-inventors during the dot-com bubble in the research triangle region. Reg. Stud. 51, 1179–1191. https://doi.org/10.1080/00343404.2016.1278300.

Miguelez, E., Raffo, J., Chacua, C., Coda-Zabetta, M., Yin, D., Lissoni, F., Tarasconi, G., 2019. Tied in: the global network of local innovation (economic research working paper no. 58), cahiers du GREThA (2007-2019). WIPO.

Motta, M., 2004. Competition policy: theory and practice. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511804038.

Nicholas, T., 2010. The role of independent invention in U.S. technological development, 1880–1930. J. Econ. Hist. 70, 57–82. https://doi.org/10.1017/S0022050710000057.

Nicholas, T., 2011a. Independent invention during the rise of the corporate economy in Britain and Japan. Econ. Hist. Rev. 64, 995–1023.

Nicholas, T., 2011b. Did R&D firms used to patent? Evidence from the first innovation surveys. J. Econ. Hist. 71, 1032–1059.

O'Donoghue, T., Scotchmer, S., Thisse, J.-F., 1998. Patent breadth, patent life, and the pace of technological progress. J. Econ. Manag. Strateg. 7, 1–32.

Ortega, J.L., 2011. Collaboration patterns in patent networks and their relationship with the transfer of technology: the case study of the CSIC patents. Scientometrics 87, 657–666. https://doi.org/10.1007/s11192-011-0363-4.

Owen-Smith, J., Powell, W.W., 2004. Knowledge networks as channels and conduits: the effects of spillovers in the Boston biotechnology community. Organ. Sci. 15, 5–21. https://doi.org/10.1287/orsc.1030.0054.

Penrose, E.T., 1951. The Economics of the International Patent System. Johns Hopkins Press, Baltimore.

Peters, L.D., Johnston, W.J., 2009. Understanding absorptive capacity from a network perspective. J. Bus. Mark. Manag. 3, 29–50. https://doi.org/10.1007/s12087-008-0036-7.

Phelps, C.C., Heidl, R., Wadhwa, A., 2012. Knowledge, networks, and knowledge networks: a review and research agenda. J. Manag. 38, 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. Int. J. Netw. Virtual Organ. 12, 352–373. https://doi.org/10.1504/IJNVO.2013.057282.

Rubens, N., Still, K., Huhtamäki, J., Russell, M.G., 2011. A network analysis of investment firms as resource routers in Chinese innovation ecosystem. J. Softw. 6, 1737–1745. https://doi.org/10.4304/jsw.6.9.1737-1745.

Sáiz, P., 2014. Did patents of introduction encourage technology transfer? Long-term evidence from the Spanish innovation system. Cliometrica J. Hist. Econ. Econom. Hist. 8, 49–78.

Sáiz, P., Zofío, J.L., 2022. The making and consolidation of the first national trademark system: the diffusion of trademarks across Spanish regions (1850-1920). Reg. Stud. 56, 256–275, 10.180/00343404.2021.1887472.

Sáiz, P., Llorens, F., Blázquez, L., Cayón, F. (Eds.), 2008. Base de Datos de Solicitudes de Patentes. España, 1878-1939. OEPM-UAM, Madrid.

Schilling, M.A., Phelps, C.C., 2007. Interfirm collaboration networks: the impact of large-scale network structure on firm innovation. Manag. Sci. 53, 1113–1126. https://doi.org/10.1287/mnsc.1060.0624.

Schmookler, J., 1966. Invention and Economic Growth. Harvard University Press, Cambridge.

Scotchmer, S., Green, J., 1990. Novelty and disclosure in patent law. RAND J. Econ. 21, 131–146.

Singh, H., Kryscynski, D., Li, X., Gopal, R., 2016. Pipes, pools, and filters: how collaboration networks affect innovative performance. Strateg. Manag. J. 37, 1649–1666. https://doi.org/10.1002/smj.2419.

Tedeschi, G., Vitali, S., Gallegati, M., 2014. The dynamic of innovation networks: a switching model on technological change. J. Evol. Econ. 24, 817–834. https://doi.org/10.1007/s00191-014-0374-4.

Ter Wal, A.L.J., 2014. The dynamics of the inventor network in German biotechnology: geographic proximity versus triadic closure. J. Econ. Geogr. 14, 589–620. https://doi.org/10.1093/jeg/lbs063.

Ter Wal, A.L.J., Boschma, R., 2009. Applying social network analysis in economic geography: framing some key analytic issues. Ann. Reg. Sci. 43, 739–756. https://doi.org/10.1007/s00168-008-0258-3.

Urzúa, C.M., 2020. A simple test for power-law behavior. Stata J. 20, 604–612. https://doi.org/10.1177/1536867X20953571.

van der Wouden, F., 2020. A history of collaboration in US invention: changing patterns of co-invention, complexity and geography. Ind. Corp. Chang. 29, 599–619. https://doi.org/10.1093/icc/dtz058.

van der Wouden, F., Rigby, D.L., 2019. Co-inventor networks and knowledge production in specialized and diversified cities. Pap. Reg. Sci. 98, 1833–1853. https://doi.org/10.1111/pirs.12432.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393, 440–442.

Whittington, K.B., Owen-Smith, J., Powell, W.W., 2009. Networks, propinquity, and innovation in knowledge-intensive industries. Adm. Sci. Q. 54, 90–122.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge.

Zacchia, P., 2018. Benefiting colleagues but not the city: localized effects from the relocation of superstar inventors. Res. Policy 47, 992–1005. https://doi.org/10.1016/j.respol.2018.03.004.