

Mejdalani, Alexandre; Gonçalves, Eduardo; Rodriguez, Rodrigo Siqueira

Article

Why cooperate? The determinants of forming inter-regional connections in the Brazilian patent network

EconomiA

Provided in Cooperation with:

The Brazilian Association of Postgraduate Programs in Economics (ANPEC), Rio de Janeiro

Suggested Citation: Mejdalani, Alexandre; Gonçalves, Eduardo; Rodriguez, Rodrigo Siqueira (2021) : Why cooperate? The determinants of forming inter-regional connections in the Brazilian patent network, EconomiA, ISSN 1517-7580, Elsevier, Amsterdam, Vol. 22, Iss. 1, pp. 71-83, <https://doi.org/10.1016/j.econ.2020.11.001>

This Version is available at:

<https://hdl.handle.net/10419/266974>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>



Why cooperate? The determinants of forming inter-regional connections in the Brazilian patent network

Alexandre Mejdalani, Eduardo Gonçalves*, Rodrigo Siqueira Rodriguez

Department of Economics, Federal University of Juiz de Fora, Rua José Lourenço Kelmer s/n, Universidade Federal de Juiz de Fora, Faculdade de Economia, Campus Universitário, Juiz de Fora – MG, Brazil

ARTICLE INFO

Article history:

Received 24 December 2017
Received in revised form 23 August 2020
Accepted 9 November 2020
Available online 1 December 2020

Keywords:

Invention networks
Probabilistic model
Co-patenting

ABSTRACT

Modern innovation activity is the product of cooperation between people, companies, and regions rather than a solitary activity performed by the traditional garage inventor. Empirical studies on cooperative innovation highlight three major types of proximities that enable tie-formation: geographical, cognitive, and social or relational proximities. Our paper aims to investigate the determinants of network formations between Brazilian regions between 2000 and 2011. We use the patent database from the National Institute of Intellectual Property (INPI) to create a network of inter-regional co-patenting in Brazil. The main results indicate that: (1) geography still plays a fundamental role in forming networks; (2) the network and its subcomponents are spatially concentrated; (3) there is an unequal geographical agglomeration pattern in the Brazilian invention system and (4) technological and relational proximity are important factors in determining new connections.

© 2021 The Authors. Production and hosting by Elsevier B.V. on behalf of National Association of Postgraduate Centers in Economics, ANPEC. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

The innovation economy seeks to study the determinants of creating and diffusing knowledge in space. While [Griliches \(1979\)](#); [Pakes and Griliches \(1984\)](#); [Jaffe \(1989\)](#); [Glaeser et al. \(1992\)](#); [Krugman \(1991\)](#) and [Jaffe et al. \(1993\)](#) contributions analyse the characteristics of geographical innovation space, a second and more recent branch investigates innovation in the relational space of connected individuals, companies and regions ([Breschi and Lissoni, 2009, 2004](#); [SINGH, 2005](#); [Maggioni et al., 2007](#); [Hoekman et al., 2009](#)).

Thus, a group of stylized facts on the behaviour of innovative activity which allows the association of geographical and relational perspectives ([Feldman and Kogler, 2010](#)) can be highlighted: (1) innovation tends to be spatially concentrated ([Jaffe, 1989](#); [Audretsch and Feldman, 1996](#)); (2) inventive activity benefits from strong connections within these agglomerations ([Cowan and Jonard, 2004](#)); (3) inter-personal innovation networks are confined to a physical space ([Singh, 2005](#)); (4) new geographical and social relations are forged from previous close contact ([Bercovitz and Feldman, 2010](#)). The challenge is a methodological construction which benefits both themes.

The methodological construction of this new analytical theme makes use of a set of instruments which overcome certain limitations of econometrics traditionally used in this field of study. Some scholars evaluate the fundamental character-

* Corresponding author.

E-mail addresses: a.mejdalani@gmail.com (A. Mejdalani), eduardo.goncalves@ufff.edu.br (E. Gonçalves), rrodriguez431@gmail.com (R.S. Rodriguez).

istics of the inventive network and investigate why individuals cooperate in short or long distances (Henderson, 2007; Bergman, 2009). Some empirical studies seek to investigate this relationship between forming a network and patent applications or citations using tools which unite classic and spatial econometrics and network analysis instruments. Although methodologically similar, these studies use different approaches, data and objectives.

With regards to the data approach, Frenken et al. (2009) and Ejerme and Karlsson's (2006) studies emphasized the regional aspects of forming ties. However, Maggioni, Nosvelli and Uberti (2007) and Crescenzi et al. (2016) use co-patenting data to investigate forming ties. Breschi and Lissoni (2005, 2004) perform an investigation on the patent citation network. From the perspective of the dimensions evaluated, Cassi and Plunket (2015) prepared a multidimensional study on French genomic patents, considering geographical, relational and technological characteristics, but without taking the specific characteristics of the nodes¹ into account. Crescenzi et al. (2016) perform a multidimensional investigation at the inventor level, with joint and specific node characteristics. The work of Breschi and Lissoni (2005 and 2004) and Ejerme and Karlsson, while it does not develop the technological dimension, it is fundamental for the methodological construction of a joint analysis of the geographical and relational dimensions.

With regard to Brazil, many studies have explored the geographical perspective of innovation, such as Gonçalves (2007), about the geographic distribution of patents, Gonçalves and Almeida (2009), about regional spillovers, and Araújo and Garcia (2019), who explored the spatial dependences of innovation in Brazilian regions, and Mascarini et al. (2019) concerning the micro-regions of the State of São Paulo. Other researchers, such as Araújo et al. (2019); Reis et al. (2018), and Sidone et al. (2016) used network topology to analyse collaboration in patents and scientific publications, respectively.

The aim of this study is to investigate the regional determinants of forming inventor ties in Brazil established through co-patenting, adopting the proximity framework. We investigate three proximity dimensions as motivators of inter-regional cooperation: distance (as a proxy of the geographic proximity), technological similarity (as a proxy of the cognitive proximity), and network centrality (as a proxy of the relational proximity)². The data used explores the relational connections formed between the regions of influence of Brazilian cities (REGICs) between 2000 and 2011, and is divided into three-year periods.

This study is divided into four parts. In the first, the nature of forming new connections between inventors is discussed, exploring the two themes in literature which explore geographical and relational space. An analysis of the network and space in the ties formed between the REGICs in Brazil is performed in the second part. Spatial analysis techniques common in spatial econometrics literature and social network graph theory are used for this purpose. A methodological framework is developed in the third part to support the empirical model to be tested. The results are presented and discussed in the fourth section. Finally, the discussion is concluded with a review of the findings in the light of the literature, and suggestions are put forward for future research.

2. Proximity dimensions in forming inventor ties

The area of literature which credits a fundamental role to relational space in forming connections between inventors who cohabit the same geographic space or otherwise attributes a fundamental factor for greater connectivity to their relative position in the network (Hoekman, Frenken and Van Oort, 2008). An agent connected to various others in the network enjoys relational advantages, since his/her field of indirect connections is wider than another individual with few connections, or another with numerous connections with very few nodes. This effect is similar within the sphere of ties between regions. It is much easier for a region connected to various others to form a new tie with another if these two have a minimum of one connection with a region in common.

However, while the presence of well-connected points in the network can increase the possibility of new ties, this can also involve forming clusters of very few well-connected regions, which act as large central nodes through which information converges and is then diffused. In their study of the connections between Swedish regions, Ejerme and Karlsson (2006) found that the formation of central nodes is highly related to a region's capacity to also generate internal nodes, which are facilitated by factors connected to urban agglomeration.

Saxenian (1996) discusses that the formation of regions favourable to development is largely due to institutional characteristics. These allow a greater connection and diffusion of knowledge between its agents. This shows the presence of social capital, constructed by various factors that bring them closer together in a network. Thus, even a highly agglomerated region has its internal ties conditioned to the presence of social capital, which allows knowledge diffusion. In contrast, Cassi and Plunket (2015) highlight that very dense networks can be detrimental to constructing ties between them and other regions in the network.

Similar to the formation of intra-regional ties, a region's capacity to form bridging ties with others has determinants associated to the way this relationship takes place and the region's relative position in relation to its neighbours. Regions

¹ "Nodes" are the basic units in network analysis. They are defined as points distributed in a relational space. The connection formed between two nodes is called a "tie".

² Other papers, such as Crescenzi et al. (2016), also include other factors that may approximate inventors, such as language and ethnicity.

Table 1
Main empirical studies which explore the determinants of forming networks.

Author(s)	Sample	Dependent Variable	Dimension	Estimation Method
Breschi and Lissoni (2005; 2010)	Italian patent citations between 1978 and 1999	Connection between patent citations	Geographical Relational	Logit
Ejermo and Karlsson (2006)	Swedish patents filed at the European Patent Office (EPO)	Affinity across regions	Geographical Relational	Weighted Least Square
Maggioni, Nosvelli and Uberti (2007)	109 European regions at NUTS 2 level, which encompasses France, Germany, Italy, Spain and the United Kingdom, during the period 1998–2002.	Co-patents between inventors	Geographical	(1) Ordinary Least Square (2) Spatial Error Model (Maximum Pseudo-likelihood)
Frenken, Hoekman, Kok, Ponds, van Oort and van Vliet (2009)	(1) Co-publication of the scientific research of 36 countries (2004) (2) Co-publication of scientific research in Europe between 1988 and 2004	Publications across regions	Geographical	Negative Binomial Regression (Gravity Model)
Cassi and Plunket (2015)	Patents recorded at the EPO between 1990 and 2006 in the areas of genetics and genomics.	Forming bridging and closure ties.	Geographical Technological Relational	Conditional Logit Multinomial Probit
Crescenzi et al. (2016)	173,180 inventors associated to patents in the United Kingdom (133,610 residents) from 1978 to 2010.	Co-patents between inventors	Geographical Organizational Cognitive Relational Ethnic and Cultural	Linear probability model: Robust POLS with fixed effects

which share neighbours (higher Jaccard Index), whose neighbours are also connected points and are central in the network (more closeness centrality) are more likely to connect than regions which share relational characteristics³.

The role of geographical distance in collaborating with knowledge production is related to growing collaboration costs with the increase in distance (Hoekman et al., 2009). Increased costs lead to the agglomeration of innovative activity (Feldman and Kogler, 2010) and, as a consequence, the formation of diverse industrial sectors beneficial to producing innovations (Feldman and Audretsch, 1999).

Bathelt et al. (2004) studied knowledge spillovers among local clusters. The literature the authors analysed is not convergent on the dominance of forming local networks on dispersed networks to create new knowledge but seems to put forward a combination between close and distant interactions. The more firms and inventors are engaged in forming inter-cluster ties (or creating pipelines), a higher amount of information enters the location and, therefore, more internal ties are formed.

The role of geography in forming these ties, therefore, has a dual direction: while forming translocal ties allows a higher local flow of information, the increased distance between innovation clusters increases connection costs, which leads to a reduction in making use of knowledge spillovers. Maggioni and Uberti (2006) highlight that inter-regional co-invention is a process involving the transmission of tacit and codified knowledge. More complex knowledge, which requires a more proximal relationship (tacit), is more dependent on the geographical function, while codified knowledge could be diffused more easily without the barrier of distance.

Regions which share similar knowledge or technological bases are cognitively close (Cassi and Plunket, 2015). Jaffe (1989) highlights that knowledge and information flow between individuals or regions more easily when there is a shared cognitive base among the nodes. Similarly, Griliches (1992) highlights that codified knowledge spillovers depend on the absorptive capacity of the firms and inventors who receive this information. From the point of view of literature on networks, absorptive capacity is associated to the ability of a node to understand the information which flows through the network, in order to identify potential collaborators from the cognition formed between overspilled and absorbed knowledge. Cassi and Plunket (2015) review two possible effects due to the cognitive or technological proximity between the nodes in literature. On one hand, agents can seek cooperation on account of affinity of knowledge, so that the closer the shared technological base, the greater the probability of connection. On the other hand, it may have a compensatory effect, if the nodes seek inventors who complement them with knowledge from other technological areas. In this case, technological similarity has a negative effect on forming ties.

Very few studies explore the relationship of probability in forming ties between inventors or regions from the viewpoint of other dimensions of analysis besides geographical. These studies present different methodologies, dimensions and forms of measurement (see Table 1). The role of geography in forming ties is evaluated in all of the studies. This is achieved by assigning a variable which measures the influence of distance in forming ties (of patents or citations) or greater affinity between two

³ These metrics are formally described in section 4, when discussed in the context of the empirical methodology used for the model of the article.

regions. The studies also consider some type of relational dimension, using relative position measurements, such as density, affinity or average level of connection. More recent research, such as [Cassi and Plunket \(2015\)](#) and [Crescenzi et al. \(2016\)](#) explore topics from an institutional perspective, considering cognitive–technological or even organizational dimensions. The latter also offers a social and economic control dimension, which includes the individual's personal characteristics.

With regards to the methodology used, probability models dominate, with the complementary use of spatial models in some cases, or methods which are used for count data distribution (Negative Binomial).

3. Spatial analysis of Co-patenting networks in Brazil

3.1. Data

This study uses data from Brazilian patents obtained from the National Institute of Intellectual Property (INPI), covering the period 2000–2011. After selecting the patent processes for useful information on the inventor's residence location, the patent data was aggregated by region to which the inventors' home municipalities belonged. Considering the inventors who collaborated during the technological production process, we compute the pairs of connections between two regions in case the inventors belong to different regions.⁴

The spatial level of analysis used in this article is based on concept of Regions of Influence of Brazilian Cities (REGIC) ([IBGE, 2013](#)). The regions of influence is a system used by the Brazilian Institute of Geography and Statistics (IBGE) to classify urban hierarchy among cities in Brazil according to social–economic influence networks, which includes services, goods, and labour dependences. The system is currently in its fifth version (2018, and previously: 1972, 1987, 1993 and 2007). As our period of study comprehends the first decade of the 20th century, we use the 2007 version. Two advantages arise from using this system. First, we avoid the bias from including small towns without any innovation infrastructure by incorporating them into a larger, as far more influential, group. Second, we do not need to worry about intraregional dependencies, such as pendulum mobility or local. Thus, our data was grouped into 161 Intermediate Areas of Urban Connectivity. This level encompasses economic agglomerations under the influence of central places. The choice of this level of regional aggregation is based on the perception that urban centers have economic relationships that coincide or not with the administrative limits of metropolises, micro–regions or states. This spatial unit of analysis fits the concept of cities that polarize a larger territory well, as conceived by [Christaller \(1966\)](#).

3.2. General network description

This article analyses 80,051 patent applications with two or more inventors, which formed 12,490 inter-regional connections (ties). Considering only unique ties, there are bidirectional 1083 ties (out of 25,760 possible connections⁵) between 124 of the 161 regions. On average, each region was connected to approximately 13 regions from this total. Among the 10 regions which were connected to the highest number of regions ([Table 2](#)), it is notable that all the polarizing cities are located in the south-southeast of Brazil, with the exception of Brasília, highlighting the State of São Paulo, which has 4 regions within the 10 most connected. If compared with the regions which formed the highest number of ties ([Table 3](#)), the same regions remained at the top, with some variations in their positions. These 10 regions single-handedly concentrate 63.4 % of the total of Brazilian inter-regional co-patenting network connections, which could indicate a tendency towards the spatial agglomeration of ties in the south-southeast region of the country.

Of the total ties, 29.9 % are links formed by regions of influence of cities in the same state, while 70.1 % are formed between those from different ones. Among the macro-regions, 54.2 % are ties between regions in the same macro-region, while 45 % are ties formed outside of it. A way of visualizing the characteristics of these ties is from the spatially distributed graph of the connection network between the regions for each three-year period.

The maps below ([Fig. 1](#)) represent the geographical distribution of all the nodes of the relational matrix. The colour of each node or city of influence indicates a higher level of connectivity, the higher the number, the darker it is. Initially some immutable characteristics of the network are noted throughout the periods: (1) the nodes with a higher level of connectivity are agglomerated in the southeast region; (2) the nodes in the north region have little or no connectivity; (3) the ties established by the cities in the north and northeast regions mainly have cities in the southeast of the country as their destination; (4) the region of influence of Brasília, the capital of Brazil, seems to be an isolated point in the central region of the country. A cluster analysis (2.2) allows a better visualization of these phenomena.

Two approaches could be used to evaluate agglomeration patterns or network dispersion: relational ([Leenders, 2002](#)) and geographical ([Anselin, 1988](#)). Firstly, from a relational viewpoint, it is possible to evaluate how close these more central regions are in relation to the network as a whole. Thus, it is possible to identify if the central regions tend to connect more between themselves or polarize other, less central regions. In the first case, the pattern observed is the agglomeration of central regions while, in the second, a dispersion pattern is noted. The second network analysis perspective is geographical. It

⁴ Similar procedure was used by [Reis et al. \(2018\)](#).

⁵ Possible connections are calculated multiplying $n(n - 1)$, where n is the number of nodes in the network.

Table 2

List of the regions (REGICs) most connected to other regions (2000–2011).

	REGIC	State	Connected Regions
1	São Paulo	SP	91
2	Rio de Janeiro	RJ	77
3	Belo Horizonte	MG	71
4	Campinas	SP	69
5	Curitiba	PR	61
6	Brasília	DF	58
7	Ribeirão Preto	SP	57
8	Florianópolis	SC	57
9	São Carlos	SP	54
10	Porto Alegre	RS	53
11	Salvador	BA	44
12	Fortaleza	CE	41
13	São José dos Campos	SP	40
13	Londrina	PR	40
13	Joinville	SC	40
13	Goiânia	GO	40
17	Recife	PE	39
18	Vitória	ES	38
19	João Pessoa	PB	34
20	Caxias do Sul	RS	32
	Brazil Average		13.45

Source: Author's Own.

Table 3

Regions (REGICs) with the most ties (2000–2011).

	REGIC	State	Ties	Average Ties/Region
1	São Paulo	SP	3986	43.80
2	Campinas	SP	2660	38.55
3	Rio de Janeiro	RJ	2512	32.62
4	Belo Horizonte	MG	1449	20.41
5	Curitiba	PR	1115	18.28
6	Porto Alegre	RS	1068	20.15
7	Ribeirão Preto	SP	802	14.07
8	Florianópolis	DF	799	14.02
9	Brasília	SC	799	13.78
10	São Carlos	SP	646	11.96
11	São José dos Campos	SP	523	13.08
12	Fortaleza	CE	466	11.37
13	Recife	PE	419	10.74
14	Salvador	BA	418	9.50
15	Manaus	AM	407	13.57
16	Joinville	SC	386	9.65
17	Araraquara	SP	328	10.58
18	Goiânia	GO	324	8.10
19	Vitória	ES	311	8.18
20	Natal	RN	273	9.10
	Brazil Average		155.15	11.53

Source: Author's Own.

is possible to identify whether the central regions agglomerate in a delimited geographical space or are dispersed throughout the country through this.

Pattern analysis is carried out using Moran's I, which provides a spatial auto-correlation coefficient for the variable of interest (Anselin, 1988), using a given connectivity criterion between the regions (spatial weight matrix). Two classes of spatial weight matrixes will be explored in this study: (i) geographical, divided into matrixes of contiguity and k-nearest neighbours and (ii) relational, based on the adjacency or contiguity of the network.

Using Leenders' methodology (2002)⁶, the binary matrix of triennial relational contiguity is used to calculate Moran's I, assuming 1 if the tie exists and 0 if it does not; regions without connections form islands. The variable analysed is the centrality of the network around the point, which indicates the importance of a point in relation to the entire network. With

⁶ Leenders (2002) argues that as the isolated points are only structurally equivalent to others in relational matrixes, they can be disregarded in the model, since they do not exert any direct or indirect relational influence on the other nodes. These relational islands were disregarded for the Moran's I relational calculation demonstrated in this study. The following procedure was adopted to construct the relational contiguity matrix: (1) data export from the network listed in Gephi software; (2) the list's compatibility with the Excel software file format, exported in ".txt" format and (3) reading the file using Geoda software.

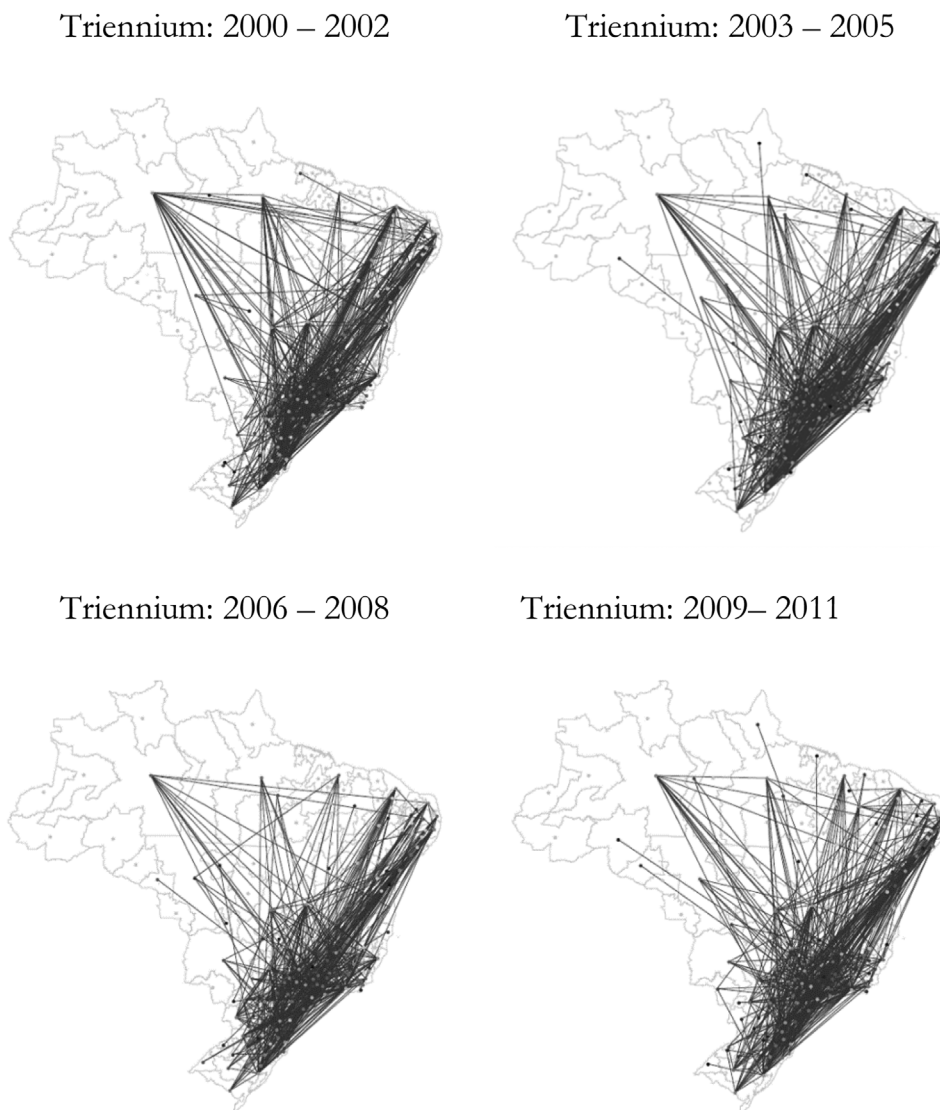


Fig. 1. Maps of Triennial Co-patenting Networks.

Source: Author's own using data from the INPI.

Table 4
Relational Moran's I for centrality by three year periods.

Triennium	Moran's I	p-value (999 p.)
2000 – 2002	0.672851	0.001 ***
2003 – 2005	0.609413	0.001 ***
2006 – 2008	0.618074	0.001 ***
2009 – 2011	0.572206	0.001 ***
2000–2011	0.296	0.001 ***

Source: Author's Own.

the relational lag of this variable, it is possible to verify if there is a tendency for the more central points in the network to connect or, in other words, a relational agglomeration in a few cities. The statistical significance was tested with 999 permutations between the network nodes (Table 4).

A tendency towards relational agglomeration around the highly central nodes in the network is observed. This indicates that highly connected nodes connect with others in this position, at least at the inter-regional level. A similar result was found by Li et al. (2015) for China.

An analysis of the geographical agglomeration pattern for the central regions in the network was carried out in order to identify if there is a spatial concentration of the nodes in a delimited physical space. Thus, the Moran's I is calculated using

Table 5
Moran's I Triennial Geographical for Centrality.

Matriz	2000 - 2002	2003 - 2005	2006 - 2008	2009 - 2011	2000 - 2011
Queen	0.002	-0.062	0.447 **	0.383 **	0.429 ***
Rook	0.002	0.062	0.441 **	0.380 **	0.424 ***
K1	0.578 **	0.346	0.273 ***	0.716 ***	0.630 ***
K2	0.410 **	0.246	0.766 ***	0.676 ***	0.648 ***
K3	0.219	0.308	0.730 ***	0.668 ***	0.640 ***
K4	0.278 *	0.394 **	0.704 ***	0.629 ***	0.630 ***
K5	0.205	0.339 *	0.631 ***	0.573 ***	0.643 ***

*10 % significance; ** 5% significance; *** 1% significance.

Note: 1) 999 permutations were carried out to calculate the pseudo-significances; 2) the regions of influence of Campina Grande (Paraíba) and Arcoverde (Pernambuco) were removed from the second triennium (2003–2005). These two cities form a closed network (only connect with each other) and are therefore spatial outliers.

Source: Author's Own.

Table 6
Percentage of Ties Formed per Triennium, 161 regions.

Y	2000–2002	2003–2005	2006–2008	2009–2011
Tie not formed (y = 0)	96.36 %	96.13 %	96.29 %	95.92 %
Tie Formed (y = 1)	3.64 %	3.87 %	3.71 %	4.08 %

Source: Author's Own.

7 weighting matrixes (queen, rook and k-nearest neighbours from 1 to 5) listed on the lines and for the 4 trienniums in the columns. Only the nodes which have a connection or, in other words, whose network centrality is greater than 0 were considered, so as not to affect the average with regions without relational weight.

The cells indicated in a dark colour identify the matrixes which have a better adjustment for the three-year periods specified or, that is, a higher significant Moran's I value (see Table 5). There is an indication of spatial agglomeration of the central cities in the network in all of the trienniums, although there are periods with higher (the last two trienniums) and lower (the first two trienniums) agglomeration. This result confirms the findings of Gonçalves (2007) and Araújo and Garcia (2019) – who evaluated the spatial dependence of Brazilian innovations.

4. Empirical model and estimation

4.1. Dependent variable and estimation

As the aim of this study is to identify the determinants of forming ties between two regions, the dependent variable comprises a pair of regions (i and j), which take on a value equal to 1 if there is a minimum of one connection between the regions and equal to zero if they do not connect during the three-year period in question. All of the possible and non-directional ties from a region are considered, which provides a total of $[n(n - 1)]/2$ potential ties to be formed in the network. As y is a binary variable, logit model estimation was selected, whose formation probability (p_{ij}) is given by (Wooldridge, 2010):

$$p_{ij} = \frac{\exp(z)}{1 + \exp(z)} \quad (1)$$

This model has its specification derived from pattern logistic distribution, the log-likelihood probability function estimation of which is achieved by maximum likelihood. This estimation method is similar to that adopted by Cassi and Plunket (2015) which, however, uses conditional logit, since the connections considered by the authors could be broken down into intra- (close connections) and extra-components (bridging ties). In this study, the interest is in evaluating the extra-component determinants, which leads to selecting a binary response model, without breaking it down.

In Table 6, the percentage of ties formed in relation to the number of potential ties in each three-year period is described. It has been noted that, on average, the regions have few connections in relation to the total potential connections. However, a more detailed analysis reveals that 49 regions were not connected to any others throughout the three-year periods, while more connected regions, as indicated in Table 2, were connected to more than 50 regions, among the 160 possible inter-regional connections.

Considering only the regions which are connected to any other in the triennium, so that isolated points are not taken into account, there are significantly higher values (Table 7). It can also be noted that the number of non-isolated nodes per triennium is not constant, at between half and two-thirds of the total regions.

As Table 8 indicates, approximately 30 % of the regions are not connected to any other during the trienniums analysed, while 34 % have a minimum of one connection in all of the three-year periods. However, 35.40 %, or 57 of the 161 regions analysed have connections in some of the trienniums but do not form connections in all of them.

Table 7
Percentage of Ties Formed per Triennium, non-isolated regions.

Y	2000–2002	2003–2005	2006–2008	2009–2011
Tie not formed ($y=0$)	88.29 %	89.06 %	89.06 %	90.91 %
Tie Formed ($y=1$)	11.71 %	10.94 %	10.94 %	9.09 %
#Regions with 1 tie or more	89	95	93	107
% of connected regions	55.3 %	59.0 %	57.8 %	66.5 %

Source: Author's Own.

Table 8
Distribution of regions by number of triennium in which a minimum of one connection is made.

# Trienniums with Ties	Number of Regions	% in Relation to the Total
Not connected	49	30.43%
Only 1 triennium	24	14.91%
Only 2 trienniums	20	12.42%
Only 3 trienniums	13	8.07%
All 4 trienniums	55	34.16%
Total	161	100%

Source: Author's Own.

Table 9
Macro-regional distribution of the regions with ($y=1$) and without connections ($y=0$) in all the trienniums.

Macro-region	#regions without any connections in any triennium	% relating to the total regions in the macro-region	#regions with connections in the 4 trienniums	% relating to the total regions in the macro-region
North	6	33.3%	3	16.7%
Northeast	29	51.8%	8	14.3%
Southeast	1	2.3%	28	63.6 %
South	6	20.0%	14	46.7%
Centre-West	7	53.8%	2	15.4%
Total	49	–	55	–

Source: Author's Own.

While the regions without connections are concentrated in the N-NE-CW, encompassing many of the areas in these regions, the S-SE macro-regions contain the regions which have a minimum of one connection in all the trienniums analysed. The southeast is highlighted, in which 63.6 % of the regions are connected to another in all of the periods (see [Table 9](#)).

5. Variables of interest

The estimation to be made aims to capture the effects of the three components below:

- (1) Geographical Distance (Geo_{ij}): measures if there is a spatial decay effect in the connection possibility. The greater the geographical distance between the two regions, the lower the tendency for connection, which should reflect in a negative sign for this coefficient. This variable is based on the geographical distance between the centroid of the polarizing cities in the region.
- (2) Relational Proximity (Rel_{ij}): is measured by the Jaccard Index, which tries to capture the effect of indirect connections or an indirect connection of the highest order (neighbours of neighbours). Thus, it is expected that the two regions which have many indirect connections are more likely to connect or, in other words, a positive sign for this coefficient. The Jaccard index is a typological metric, based on neighbourhoods used to predict connections between two nodes in the network, according to [Wasserman and Faust \(1995\)](#) and [Jackson \(2008\)](#). This article uses the index to measure the relative proximity of two points. It considers the ratio between the number of common neighbours between two points i and j ($I_i \cap I_j$) and the total neighbours for these ($I_i \cup I_j$):

$$J(i, j) = \frac{|I_i \cap I_j|}{|I_i \cup I_j|} \quad (2)$$

- (3) Technological Proximity (Tec_{ij}): attempts to measure if the technological similarity between two regions increases their chances of forming a new connection. The Jaffe Index (1986) is used, which is constructed with the patent classes to measure the technological proximity between two regions. The higher the level of patenting in the same technological classes, the higher the index, which indicates technological similarity. This proximity can have two types of effect: (i) technologically closer regions have a greater cognitive affinity, which facilitates forming ties; (ii) regions can seek others

Table 10
Descriptive statistics for the control variables.

Variable	Average	Standard Deviation	Min	Max
Difference in GDP per capita	0.7554	0.5508	0	3.0764
Difference in Centrality	0.2024	0.1749	0	0.6931
Difference in R&D Capacity	0.0030	0.0032	0	0.0221
Same state	0.0523	0.2227	0	1

Note: The differences are defined as the absolute difference in the natural logarithms for the pair (i and j).

Source: Author's Own.

with a less technological similarity, in order to achieve technological complementarity. Thus, the coefficient sign could be positive or negative.

6. Control variables

In order to control the effect that other factors may have on the tendency to form ties, the following were included: (i) level of economic development in the region; (ii) relative position of the node in the network and (iii) regional effort in research activities.

- i GDP per capita (GDP_{ij}): attempts to capture if the connections tend to be formed between regions with a similar level of development (Maggioni et al., 2007; Maggioni and Uberti, 2007). The GDP per capita for the regions is a measure constructed based on Gross Domestic Product data calculated by the Brazilian Institute of Geography and Statistics (IBGE), National Household Sample Survey (PNAD) population data and Demographic Censuses from 2000 and 2010. The measurement used here refers to the absolute difference between the GDP per capita natural logarithm for regions i and j , which form potential ties.
- ii Degree of centrality ($CENT_{ij}$): attempts to capture the relative position of the node in the network or, in other words, if the node is a central region and has a score with ranges from 0 to 100. The variable is constructed by the absolute difference between the natural logarithm for relative centrality for regions i and j . It is expected that the regions with a close centrality connect more easily (Crescenzi et al., 2016). Thus, a positive coefficient is expected for this variable. Based on Wasserman and Faust (1995) and Jackson (2008), this article uses the Closeness Centrality metric, which measures how easy it is for one node to connect to others within the network. The number of connections with a shorter path l between a node i and the other j nodes in the network is used as a proximity measure:

$$CC_i = \frac{(n-1)}{\sum_{i \neq j} l_{ij}} \quad (3)$$

- iii Capacity to carry out research and development ($R\&D_{ij}$): the variable is constructed using the absolute difference between the natural logarithm of workers in technical-scientific occupations (PoTec), comprising biotechnology and biomedical professionals, engineers, researchers, Maths and statistics professionals, computer system analysts, physicists, chemists, space and atmosphere professionals and architects from regions i and j in the regions.⁷ The variable was constructed with data taken from the Ministry of Labour and Employment (MTE) Annual Social Information Report, in the absence of regionalized data on R&D expenditure. Ejermeo and Karlsson (2005) put forward two possible effects for the R&D variable. On one hand, regions with a high research capacity make more connections with other regions. On the other, these regions can be self-sufficient in research, having fewer incentives to connect with others. A positive sign is expected for this coefficient in one case and a negative sign in the other.

In addition, a dummy case is included of the polarizing cities for regions i and j belonging to the same state, which takes on the value of 1. The inclusion of this variable is an attempt to capture a possible preference for inventors to connect to other regions in the same state. Table 10 presents the descriptive control statistics.

7. Empirical model

The model is designed to estimate the relation for the three dimensions (geographical, relational and technological), with the probability of forming new ties and including other social and economic controls. The variables of interest are in

⁷ According to Araújo et al. (2009) and Gusso (2006) there is a 90% correlation between firms' external and internal R&D expenditure and PoTec, which indicates that this variable is an appropriate proxy for private R&D capacity.

Table 11
Determinants of tie formation in Brazilian inventive regions by triennials.

Variables	2000–2002		2003–2005		2006–2008		2009–2011	
	Logit	WLogit	Logit	WLogit	Logit	WLogit	Logit	WLogit
Geographical Distance	−0.307*** (0.0979)	−0.257** (0.112)	−0.458*** (0.0878)	−0.486*** (0.104)	−0.371*** (0.0838)	−0.350*** (0.103)	−0.422*** (0.0804)	−0.436*** (0.0943)
Technological Proximity	6.502*** (0.607)	6.545*** (0.611)	5.826*** (0.565)	5.741*** (0.570)	4.874*** (0.492)	4.725*** (0.499)	5.663*** (0.474)	5.649*** (0.477)
Relational Proximity	3.875*** (0.176)	3.847*** (0.178)	4.484*** (0.217)	4.418*** (0.218)	3.821*** (0.195)	3.754*** (0.197)	3.444*** (0.192)	3.420*** (0.193)
Difference in GDP per capita	−0.218 (0.176)	−0.206 (0.189)	−0.213 (0.165)	−0.351* (0.186)	−0.0200 (0.166)	−0.191 (0.179)	0.105 (0.149)	0.105 (0.164)
Difference in Centrality	−3.541*** (0.897)	−3.583*** (0.904)	−2.171*** (0.811)	−2.177*** (0.812)	−3.130*** (0.797)	−3.220*** (0.815)	−2.993*** (0.769)	−3.045*** (0.772)
Difference in R&D Capacity	0.573*** (0.0532)	0.561*** (0.0538)	0.633*** (0.0504)	0.612*** (0.0512)	0.509*** (0.0471)	0.477*** (0.0479)	0.542*** (0.0431)	0.535*** (0.0441)
Same state	0.614*** (0.224)	0.665*** (0.227)	0.567*** (0.208)	0.568*** (0.211)	0.686*** (0.197)	0.719*** (0.200)	0.842*** (0.179)	0.852*** (0.181)
W.Difference in GDP per capita		−0.294** (0.136)		−0.0954 (0.196)		−0.0341 (0.186)		0.846 (0.649)
W.Difference in Centrality		0.478 (0.690)		0.753 (0.626)		−0.401 (0.715)		−0.144 (0.182)
W.Difference in # of researchers		75.70** (35.61)		100.1*** (37.42)		159.7*** (34.03)		27.30 (29.90)
Constant	−6.399*** (0.750)	−6.766*** (0.813)	−5.516*** (0.669)	−5.398*** (0.736)	−5.127*** (0.642)	−5.297*** (0.724)	−5.023*** (0.606)	−4.996*** (0.658)
Observations	12,880	12,880	12,880	12,880	12,880	12,880	12,880	12,880

Standard errors between brackets. *** p < 0.01, ** p < 0.05, * p < 0.1.

A spatial weights matrix of k = 3 closest neighbours was used, from the centroid of the central REGIC city.

Source: Author's Own.

the natural logarithm and the control variables are with the difference in the natural logarithms for the variable in *i* and *j*. Therefore, it can be written as:

$$\begin{aligned}
 P(y = 1x) = & \beta_0 + \beta_1 \ln(Geo_{ij}) + \beta_2 \ln(Rel_{ij}) + \beta_3 \ln(Tec_{ij}) \\
 & + \beta_4 [|\ln(GDP_i) - \ln(GDP_j)|] + \beta_5 [|\ln(CENT_i) - \ln(CENT_j)|] \\
 & + \beta_6 [|\ln(R\&D_i) - \ln(R\&D_j)|] + \beta_7 State + \varepsilon
 \end{aligned} \tag{4}$$

In which *Geo_{ij}* indicates the geodesic distance between *i* and *j*; *Rel_{ij}*, the Jaccard Index between *i* and *j*, and *Tec_{ij}*, the Jaffe Index calculated between *i* and *j*. For the controls, GDP indicates the gross domestic product per capita in the region specified; CENT, indicates the region's relative position (node) in the network; R&D, the proportion of workers employed in technical activities and state indicates if *i* and *j* are polarized regions by cities within the same state when equal to 1, and different states if equal to 0.

A second model to be tested seeks to capture the effects of the spatial neighbours in forming the tie (Equation 7). A spatial lag model (SAR) is adopted for this, where the control variables, GDP per capita, and PoTec are spatially lagged and the absolute difference in the natural logarithms is extracted. It is intended to verify if there is an influence from the neighbourhood to establish inter-regional ties. It is expected that two regions with rich surroundings or a high number of researchers (small absolute difference) connect less since their inter-regional connections will benefit their neighbours. However, regions with poorer surroundings or very few researchers will connect more with regions whose surroundings present advantages that meet their needs.

$$\begin{aligned}
 P(y = 1x) = & \beta_0 + \beta_1 \ln(Geo_{ij}) + \beta_2 \ln(Rel_{ij}) + \beta_3 \ln(Tec_{ij}) \\
 & + \beta_4 [|\ln(GDP_i) - \ln(GDP_j)|] + \beta_5 [|\ln(CENT_i) - \ln(CENT_j)|] \\
 & + \beta_6 [|\ln(R\&D_i) - \ln(R\&D_j)|] + \beta_7 State \\
 & + \beta_8 W [|\ln(GDP_i) - \ln(GDP_j)|] + \beta_9 W [|\ln(CENT_i) - \ln(CENT_j)|] \\
 & + \beta_{10} W [|\ln(GDP_i) - \ln(GDP_j)|] + \varepsilon
 \end{aligned} \tag{5}$$

8. Results

The estimation results are shown in [Table 11](#), separated into trienniums. The results for spatial (identified by the prefix W) and non-spatial regressions are also included. Analysing the regression coefficient values, the three variables of interest (Geo_{ij} , Rel_{ij} and Tec_{ij}) present a statistical significance and have an expected sign, both for the triennial regressions and that for the entire period.

The variable (Geo_{ij}) indicates that the greater the distance between the regions, the less probability of connection between two regions, as expected. This result indicates the same effect observed in the empirical literature analysed (section 2) that space still has a major influence in forming networks in Brazil. It also corroborates the findings of [Frenken et al. \(2009\)](#), which show that the “space of places” continues to be a major determinant of cooperation, even with the decreasing costs of new information and communication technologies in the period. If the “death of distance” were happening, we would expect the coefficient of the geographical distance to become closer to zero in the latest periods. However, we do not observe this. The coefficient fluctuates between -0.30 and -0.46, indicating that the role of distance is not becoming less important in Brazil.

Concerning the second variable of interest (Rel_{ij}), which represents the relational dimension, the coefficient has a positive sign in the triennial regressions and also the entire period, revealing that regions connected to various others are more likely to form ties. [Cassi and Plunket \(2015\)](#) find a positive result for the number of common partners variable between French inventors and the rest of Europe. [Breschi and Lissoni \(2004\)](#) use degrees of geodesic distance⁸, and find that the closer nodes in the network are more connected. The model estimated by [Ejermeo and Karlsson \(2006\)](#) includes the relational dimension as a dependent variable and finds that the sign for the number of patents independent variable is positive in Sweden. [Crescenzi et al. \(2016\)](#) research has a closer parallel to the Rel_{ij} variable and finds positive values for social proximity in the United Kingdom.

The coefficients obtained for the technological dimension of proximity (Tec_{ij}) show that technologically similar regions are more likely to connect. In this case, cognitive affinity tends to prevail because it may facilitate communication and learning with inventors. This effect was also observed by [Cassi and Plunket \(2015\)](#) in France and [Crescenzi et al. \(2016\)](#) for the United Kingdom.

The results found for the variables of interest in Brazil are in line with those in the literature evaluated throughout the article: (1) there is a spatial weight decay as the distance increases between the nodes ([Hoekman et al., 2009](#); [Henderson, 2007](#)); (2) regions with more indirect connections tend to connect more easily and form network clusters ([Bathelt et al., 2004](#); [Ejermeo and Karlsson, 2005](#)) and (3) technological proximity has a significant influence on connecting two regions ([Cassi and Plunket, 2015](#); [Crescenzi et al., 2016](#)).

The coefficient for the difference in GDP per capita, although with the same sign in the first three trienniums, is not significant. This result may indicate that the effect of the GDP per capita on the probability of connections between regions is intertemporal and has a maturation period greater than that observed by the inventors at the time the tie is formed. The coefficients obtained for the difference in centrality and R&D capacity are positive. This shows that there may be a compensatory effect in these variables. More peripheral regions have higher connections with the more central ones, and regions with less R&D capacity connect with those with a higher capacity, indicating possible R&D spillovers between regions with different appropriations of professionals in technological areas. The centrality coefficient has a negative sign in this estimation, in contrast to the positive sign for the entire period. This sign indicates the opposite effect: central regions tend to connect with other central regions in the short-term, while less central regions seek more central ones to form ties over longer periods. The coefficient for the difference in R&D capacity had the same effect observed for the entire period. This result is similar to that found by [Ejermeo and Karlsson \(2005\)](#) who highlighted the fact that regions with very few researchers tended to connect with regions with a higher R&D capacity, in order to overcome regional R&D shortcomings.

The results for the spatial logit presented signs along the same lines and with similar values, leading to comparable interpretations on the variables of interest and control throughout the trienniums, although some coefficients lost their significance. In this model, spatial lags were included in the controls (GDP per capita, centrality and R&D capacity), in order to capture the effect of the geographical neighbourhood. The result of the R&D capacity coefficient is positive and significant in three of the four estimations. In light of this, when the average spatial difference between the number of R&D professionals is high, ties are more likely to be formed. On the other hand, two regions whose neighbours are home to many researchers, which corresponds to a lower difference in the variable log, have little likelihood of connecting. This result also suggests the importance of the geographical surroundings in making invention ties.

9. Conclusions

This article sought evidence of the influence of three dimensions of proximity (geographical, technological and relational) on the tendency to form ties between inventors in different regions, controlling other factors which may influence forming ties, such as differences in levels of GDP per capita, centralities, R&D capacity and location in the same state. The empirical methodology uses instruments suggested by spatial econometrics, with the inclusion of metrics adopted by network analysis graph theory, such as notions of relative position, forming relational clusters, and organizing data by adjacency matrixes.

⁸ Shortest path which unites two nodes in the network.

Our main results indicate that inter-regional co-patenting can be explained by a combination of causalities that exceeds the influence of geographical distance. By using a social network approach, we verified that central nodes play an important role in approximating distant regions with a few to no direct connections. The nine most connected regions in the network are located in the southern part of Brazil, and the four most connected cities (São Paulo, Campinas, Rio de Janeiro, and Belo Horizonte) are only a few hundred kilometres apart, forming a highly agglomerative cluster in the Southeast Region of Brazil.

The role of the node's position in the network and technological proximity between regions with a potential connection is fundamental to determine the formation of new links, although geographical distance also plays an important role in connections. The existence of the tendency to form spatial innovation clusters within which the regions connect more intensively is also observed.

Our results have an important policy implication as they identify which regions have a national, regional, and a peripheral role in the Brazilian network of co-patents. This information can be useful to focus investment in infrastructure, human resources, and capabilities in regions that may spill over their benefits to its peers.

This article is subject to limitations and possible extensions. Firstly, only bridging ties between regions were analysed. The relationship between the cities in each region, which could indicate a proximal tie, was not included and could be a future advance. Secondly, the construction of the technological matrix only adopted the Jaffe approach for patent classes, while in the literature there are other methods to calculate technological distance, which could have their potential explored. And third, our database can only capture cooperation through co-patenting. Therefore, research and development cooperation that does not lead to a patent but builds capabilities is not considered.

Funding

This work was supported by the National Scientific and Technological Development Council (CNPq) and Foundation to Support Research for the State of Minas Gerais (FAPEMIG).

Uncited references

Anselin (1995), Araújo et al., 2009, Breschi and Lissoni (2010).

Acknowledgements

The authors gratefully acknowledge the contribution of the research team from the National Institute of Intellectual Property (INPI) by the patent database.

References

- Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Springer Science & Business Media.
- Anselin, L., 1995. *Local indicators of spatial association – LISA*. *Geogr. Anal.* 27 (2), 93–115.
- Araújo, V.D.C., Garcia, R., 2019. Determinants and spatial dependence of innovation in Brazilian regions: evidence from a Spatial Tobit Model. *Nova Econ.* 29 (2), 375–400.
- Araújo, B., Cavalcante, L., Alves, P., 2009. *Variáveis proxy para os gastos empresariais em inovação com base no pessoal ocupado técnico-científico disponível na Relação Anual de Informações Sociais (RAIS)*. Radar: tecnologia, produção e comércio exterior, 5., pp. 16–21.
- 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 (3–4), 235–280.
- Audretsch, D., Feldman, M., 1996. *R&D spillovers and the geography of innovation and production*. *Am. Econ. Rev.*, 630–640.
- Bathelt, H., MalMBERG, A., Maskell, P., 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. In: *Progress in Human Geography (Vol. 28)*, pp. 31–56.
- Bercovitz, J., Feldman, M., 2010. The mechanisms of collaboration in inventive teams: composition, social networks, and geography. *Res. Policy* 40 (10), 81–93.
- Bergman, E., 2009. *Embedding network analysis in spatial studies of innovation*. *Ann. Reg. Sci.* 43 (559–565), 2009.
- Breschi, S., Lissoni, F., 2005. *Cross-Firm Inventors and Social Networks: Localized Knowledge Spillovers Revisited*. *Annales d'économie et de statistique*, no. 79/80.
- Breschi, S., Lissoni, F., 2009. *Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows*. *J. Econ. Geogr.*, lbp008.
- Breschi, S., Lissoni, F., 2010. *Knowledge Networks From Patent Data: Methodological Issues and Research Targets*. KITeS, Centre for Knowledge, Internationalization and Technology Studies, Università Bocconi, Milano, Italy.
- Cassi, L., Plunket, A., 2015. *Research collaboration in Co-inventor networks: combining closure, bridging and proximities*. *Reg. Stud.* 49 (6), 936–954.
- Christaller, W., 1966. *Central Places in Southern Germany*. Prentice-Hall/Englewood Cliffs.
- Cowan, R., Jonard, N., 2004. *Network structure and the diffusion of knowledge*. *J. Econ. Dyn. Control* 28, 1557–1575.
- Crescenzi, R., Nathan, M., Rodríguez-Pose, A., 2016. *Do inventors talk to strangers? On proximity and collaborative knowledge creation*. *Res. Policy* 45 (1), 177–194.
- Ejermo, O., Karlsson, C., 2006. *Interregional inventor networks as studied by patent coinventorships*. *Research Policy*, Elsevier 35 (3), 412–430.
- Feldman, M., Kogler, D., 2010. *Stylized facts in the geography of innovation*. In: Hall, B.H., Rosenberg, N. (Eds.), *Handbook of Economics of Innovation*. Elsevier, cap. 8.
- Frenken, K., Hoekman, J., Kok, S., Ponds, R., van Oort, F., van Vliet, J., 2009. *Death of Distance in Science? A Gravity Approach to Research Collaboration*. *Innovation Networks, Understanding Complex Systems*, pp. 43–57.
- Glaeser, E., Kallal, H., Scheinkman, J., Shleifer, A., 1992. *Growth in cities*. *J. Polit. Econ.* 100, 1126–1152.
- Gonçalves, E., 2007. *O padrão espacial da atividade inovadora brasileira: uma análise exploratória*. *Estud. Econômicos* 37 (2), 405–433.
- Gonçalves, E., Almeida, E., 2009. *Innovation and spatial knowledge spillovers: evidence from Brazilian patent data*. *Reg. Stud.* 43 (4), 513–528.
- Griliches, Z., 1979. *Issues in assessing the contribution of research and development to productivity growth*. *Bell. J. Econ.*, 92–116.
- Gusso, D., 2006. *Agentes Da Inovação: Quem Os Forma, Quem Os Emprega*. Tecnologia, exportação e emprego, Brasília: Ipea, pp. 397–444.

- Henderson, J., 2007. [Understanding knowledge spillovers](#). *Reg. Sci. Urban Econ.* 37 (n. 4), 497–508.
- Hoekman, J., Frenken, K., van Oort, F., 2009. [The geography of collaborative knowledge production in Europe](#). *Ann. Reg. Sci.* 43 (n. 3), 721–738.
- IBGE - Instituto Brasileiro de Geografia e Estatística, 2013. *Divisão Urbano-regional*. Rio De Janeiro.
- Jackson, M., 2008. [Social and Economic Networks](#). Princeton University Press.
- Jaffe, A., 1989. [Real effects of academic research](#). *Am. Econ. Rev.*, 957–970.
- Jaffe, A., Trajtenberg, M., Henderson, R., 1993. [Geographic localization of knowledge spillovers as evidenced by patent citations](#). *Q. J. Econ.* 434, 578–598.
- Krugman, P., 1991. [Geography and Trade](#). MIT press.
- Leenders, R., 2002. [Modeling social influence through network autocorrelation: constructing the weight matrix](#). *Soc. Networks* 24, 21–47.
- Li, D., Wei, Y., Wang, T., 2015. [Spatial and temporal evolution of urban innovation network in China](#). *Habitat Int.* 49, 484–496.
- Maggioni, M., Uberti, E., 2006. [International Networks of Knowledge Flows: an Econometric Analysis](#). *Papers on Economics and Evolution 2005-2019*, Philipps University Marburg. Department of Geography.
- Maggioni, M., Nosvelli, M., Uberti, E., 2007. [Space vs. Networks in the geography of innovation: a european analysis](#). *Papers in Regional Science* 86 (3), 471–494.
- Mascarini, S., Garcia, R.C., Roselino, J.E., 2019. [Analysis of the effect of territorial factors on regional innovation in the state of São Paulo, Brazil](#). *Revista Brasileira de Estudos Regionais e Urbanos* 13, 183–200.
- Pakes, A., Griliches, Z., 1984. [Patent and R&D at the Firm Level](#). *R&D, Patents, and Productivity*, University of Chicago Press, pp. 55–72.
- Reis, R.C., Gonçalves, E., Taveira, J.G., 2018. [Determinants of inventive collaborations in Brazilian interregional and international networks](#). *Rev. Bras. Inovaãõ* 17, 287–316.
- Sidone, O.J., Haddad, E.A., Mena-Chalco, J.P., 2016. [Science in Brazilian regions: development of scholarly production and research collaboration networks](#). *Transinformao* 28 (1), 15–31.
- Singh, J., 2005. [Collaborative networks as determinants of knowledge diffusion patterns](#). *Manage. Sci.* 51 (5), 756–770.
- Wasserman, S., Faust, K., 1995. [Social Network Analysis: Methods and Applications](#). *Structure Analysis in Social Sciences*, Book 8. Cambridge University Press.
- Wooldridge, J., 2010. [Econometric Analysis of Cross Section and Panel Data](#). MIT press.