Revista Brasileira de Inovação
ISSN 2178-2822
DOI: https://doi.org/10.20396/rbi.v20i00.8661402

ARTICLE

An expanded knowledge production function: evidence from Brazil with a dynamic spatial panel approach

* *Pedro Henrique Batista de Barros*
Universidade de São Paulo, São Paulo (SP), Brasil
E-mail: pedrohbarros@usp.br

** *Adirson Maciel de Freitas Júnior*
Universidade de São Paulo, São Paulo (SP), Brasil
E-mail: adirson@usp.br

RECEIVED: 27 SEPTEMBER 2020 REVISED VERSION: 06 APRIL 2021 ACCEPTED: 09 JUNE 2021

ABSTRACT

This paper uses a theoretical motivation for an Expanded Knowledge Production Function (EKPF) that encompasses both path dependence and spatial spillovers to search for evidences in Brazil using a Dynamic Spatial Panel Data approach. The purpose is to identify the determinants of knowledge production in the 2005-2015 period as well as its temporal evolution, using innovation patents as proxies. Regarding its spatial distribution, we identified a North-South disparity for the knowledge production in Brazil, with Southeast and South producing a large part of the country's patents. Based on the EKPF, we confirmed the importance of path dependence and knowledge spillovers to explain the Brazilian innovation. In addition, population density, which generates Jacobian externalities and economies of agglomeration, is an important structural feature in the short run while the number of researchers in universities and an increased economic scale are essential to knowledge production in the long run.

KEYWORDS | Knowledge Production Function (KPF); Path Dependence; Spatial Spillover; Dynamic Spatial Panel.

1. Introduction

The economic development of a country is associated, among other factors, to its ability to generate scientific and technological knowledge, as they are likely to become innovations, which increases productivity and competitiveness of the economy. The support of technological innovation as a driving force of economic growth dates back to Solow (1956) and the emergence of long-term development theories. Despite showing its importance, Solow (1956) treated technology as exogenous. Subsequently, Romer (1990) worked this limitation, making technological innovation an endogenous variable with a key role in the process of long-term economic growth. The creation of ideas can generate new combinations of inputs, better goods and services, increasing the society general well-being. Therefore, knowledge creation and technological progress, by intensifying productivity, is an economic growth booster.

However, the creation of knowledge and innovations often entails high costs and, since they are non-rival goods, the returns obtained with the invention are usually insufficient to generate incentives for their full development. In this context, Douglas North (1981) argues that the development of intellectual property rights, the legal basis for the patent system, is a major contributor to modern economic growth. Therefore, the consolidation of institutions that guarantee intellectual property rights in a given country is an essential condition to economic development. The monopoly power induced by intellectual property makes it possible to increase the return on investment in innovations, thus encouraging the generation of knowledge.

In this context, the present paper aims to contribute to the literature on technological development and knowledge creation by designing an empirical evaluation of the determinants of knowledge production in Brazil at the microregion-level, with special focus on the process of path dependence and spatial spillover, based on theoretical and empirical contributions of Griliches (1979), Jaffe (1989), Fischer, Scherrngell and Reismann (2009) and Autant-Bernard and LeSage (2011). In particular, we propose to use the Dynamic Spatial Panel method developed by Yu, De Jongk and Lee (2008), which is capable of considering both spatial spillovers and path dependence in a same empirical design. Thus, this empirical approach allows testing the direct, indirect and total effects induced by spatial spillovers and, at the same time, to access the long run effects generated from path dependence.

To capture innovation and knowledge production, the number of patents created by a particular country is a proxy widely used in the literature (GRILICHES, 1990; ALBUQUERQUE *et al.*, 2002; GONÇALVES, 2007; GONÇALVES and ALMEIDA, 2009; FREITAS; GONÇALVES; MONTENEGRO, 2010; MONTENEGRO; GONÇALVES; ALMEIDA, 2011; GONÇALVES; MATOS; ARAÚJO, 2018; GONÇALVES; OLIVEIRA; ALMEIDA, 2020). According to Miranda and Zucoloto (2015), patents are an important indicator of the presence of "knowledge with an innovative profile". Therefore, we propose to use the number of patents[1] in the microregions to estimate an Expanded Knowledge Production Function in Brazil for the 2005-2015 period.

This article is structured in four more sections besides this introduction. The second deals with the theoretical proposition of an Expanded Knowledge Production Function, as well as an investigation in the literature on the determinants of knowledge production. In the third section, we present the empirical approach and the database used. The results found and their analysis are performed in the fourth section. Finally, the fifth section presents the final considerations.

2. Theoretical Framework

The Knowledge Production Function (KPF) was first formulated and empirically tested in Griliches (1979; 1984). The model proposed is defined as

$$Y = F(X, K, u) \tag{1}$$

where F ( ) is the production function of firms that relates the product, Y, to the inputs X, K e u and, where X is a vector of the capital and labor inputs, K is a measure of the current state of scientific and technological knowledge, which to a certain extent is determined by past research and development (R&D); captures all other factors that influence the level of productivity and production of knowledge. Griliches (1979) defines the F ( ) with a Cobb-Douglas functional form as

$$Y = DC^{\alpha} L^{\beta} K^{\gamma} e^{\lambda t + u} \tag{2}$$

where D is a constant; C and L are capital and labor respectively; t is a time indicator; e is the basis of the natural logarithm; and α, β, γ are the parameters to be estimated empirically.

1 It is worth mentioning that we used patents based on the first depositor's residence, which differs from part of the literature that uses the inventor's residence. This choice reflects the discontinuation of the BADEPI (*Base de Dados Estatísticos sobre Propriedade Intelectual*) database that contained patents by inventor's residence only up to 2012.

In a later formulation, Griliches (1984) uses the number of patents created as an indicator of knowledge production, defining a reduced form patenting model

$$p_{i,t} = dt + \beta\dot{k}_{i,t} + v_{i,t} \tag{3}$$

in which β is the elasticity of patents in relation to an increase of knowledge; d is a measure of temporal propensity to create patents; v is the part of the patents that cannot be explained by the increment of knowledge or temporal tendency.

Later, in a theoretical advance, Jaffe (1989) attempt to measure effects of spillovers in knowledge creation by considering geographic units, defined as

$$\log(P_{ikt}) = \beta_{1k}\log(I_{ikt}) + \beta_{2k}\log(U_{ikt}) + \beta_{3k}[\log(U_{ikt})\log(C_{ikt})] + \epsilon_{ikt} \tag{4}$$

where i refers to the observation unit (municipalities, microregions, states); k is the technological area; t is a time index; P is the number of patents; I is the R&D carried out by the industries; U is the R&D undertaken by universities. Lastly, C is a geographic measure that seeks to capture spillovers by considering the coincidence between activities of university and industrial research. On the other hand, Fischer *et al.* (2009) proposed

$$Q_{it} = A_{it}\, g(L_{it}, C_{it}) \tag{5}$$

where $i = 1, \ldots, N$ are the regions; while $t = 1, \ldots, T$ is a time denotation; L and C are the factors of production, labor and capital, respectively. $g(\ldots)$ is a homogeneous function of degree one with decreasing returns of scale for capital inputs (C) and labor (L). Lastly, Q is the value of production and A is an index of technological efficiency defined as $A_{it} = A(K_{it}, K_{it}^*)$ being that K are K^* the knowledge stocks inside and outside the region.

In addition, Fischer, Scherrngell and Reismann (2009) argue that it is possible that a portion of the knowledge created to be appropriated by neighboring regions due to social networks, seminars and scientific conferences, legal transfers, reverse engineering, etc. Then, due to the occurrence of knowledge spillover in neighboring regions, the variable K_{it}^* should be considered taking into account this non-complete appropriation. The authors also defend the need to include the time lag of the knowledge spillovers of the neighboring regions because its effects often do not occur in the same period. Therefore, the spillover effect must be temporally and spatially lagged in a Cobb-Douglas estimation.

Autant-Bernard and LeSage (2011) developed a theoretical model for KPF that incorporates not only the knowledge spillover generated in a given region, but also the externality of its own inputs and determinants. Considering that the spatial dependence is common in the literature[2], the authors incorporate this spatial dependency explicitly,

$$\boldsymbol{I} = \psi \boldsymbol{W}\boldsymbol{I} + \boldsymbol{r}\beta_1 + \boldsymbol{W}\boldsymbol{r}\beta_2 + \varepsilon \tag{6}$$

where $\boldsymbol{I}$ represents a (logarithmic) vector of observations from innovations performed by the n regions; $\boldsymbol{r}$ is a vector representing the measurable inputs of the knowledge production; $\boldsymbol{W}$ is a $n \times n$ spatial lag matrix that seeks to capture the structural configuration between the regions; ψ captures the force of spatial dependence and β_2 the spatial spillovers from $\boldsymbol{r}$.

Next, we seek to highlight the literature about the main variables of the knowledge production function to guide us in the empirical estimations. Lucas (1988) and Romer (1990), in the endogenous growth theory, have shown that innovation is among the "deep" causes of economic progress. This is due to the existence of positive externalities of knowledge; and its generation occurs mainly through expenditures on Research and Development (R&D), whether public or private. Freeman (1988) emphasizes the importance of articulation between the educational system and the productive sector in the generation of knowledge, especially through universities that supply skilled labor to firms, besides the basic and applied research that contribute directly. In this context, internal R&D performed in private companies, combined with those performed in universities and research institutes are key elements in the production of knowledge.

However, the production of knowledge and innovation does not take place in an isolated and independent way, since it reflects attitudes and paths taken previously in a historical construction (NELSON, 1996). According to Arthur (1989), technological development tends to assume a pre-established trajectory, in a path-dependence phenomenon. Hence, the amount of innovation created in $t - 1$ is an important determinant of knowledge produced in period t. For

2 For example: Autant-Bernard (2001), Autant-Bernard, Mairesse and Massard (2007), Parent and LeSage (2008). For Brazil, spatial dependence has also been a recurrent process: Gonçalves (2007), Gonçalves and Almeida (2009), Montenegro, Gonçalves and Almeida (2011), Araújo and Garcia (2019).

that reason, Ejermo (2005) emphasizes the need to include the temporally lagged dependent variable in the knowledge production function since initial advantages help to determine the future development of knowledge, leading, in the long run, to increasing inequalities between regions.

Krugman (1991) argues that scientific and technological knowledge, as a source of increasing returns to scale, can unleash attractive forces for similar activities, resulting in a process of geographical concentration. Thus, the spatial location is also an important factor in explaining the differences in economic growth rate, as well as in the development and dissemination of innovations. Jacobs (1969) supports the importance of agglomeration and geographic proximity, which make it possible to increase the diffusion, transmission and exchange of knowledge. This is linked to urbanization, which leads to the concentration of economic activities known as agglomeration economies. From this perspective, Griliches (1992) emphasizes the importance of knowledge spillovers related to geographical concentration and urbanization, which result from positive externalities inherent to the innovative process. These spillovers only occur when there is geographic proximity, and can induce the increase of the spatial concentration of this activity.

In relation to the analysis of the spatial distribution of the innovative activity in Brazil, Gonçalves (2007) identified the existence of spatial autocorrelation in the production of knowledge, with regions with high level of technological activity having neighbors with similar characteristics. In addition, the author found a North-South spatial pattern for the Brazilian innovation distribution, with the southeast and south regions having a high technological activity in contrast to the low values for Northeast, North and Central West. Indeed, Rodriguez and Gonçalves (2017) confirmed a positive spatial association for technological innovation in the Southeast and South regions of Brazil, particularly in the state capitals and Campinas region.

Freitas, Gonçalves and Montenegro (2010) sought to investigate the technological inequality between Brazilian states from 1990 to 2001 and also found evidence of spatial concentration of innovation, with the existence of significant spatial clusters, especially in the Southeast and South. They also identified evidence of a convergence process between Brazilian states, with undeveloped regions presenting a higher growth rate than more consolidated ones. Similar results were obtained by Oliveira, Gonçalves and Almeida (2016), for Brazil, and Barros *et al.* (2019) and Freitas Júnior *et al.* (2021), for the Southern region of the country, that also identified a spatial concentration of the innovative activity, indicating the importance of the spatial component, besides the occurrence of a catching up process between the regions.

For the Knowledge Production Function, we highlight Gonçalves and Almeida (2009) that estimate it for the Brazilian microregions in the year 2000 with a Spatial Autoregressive Model (SAR) and found that knowledge spillover is an important determinant of Brazilian innovation. In addition, factors such as R&D performed by universities and companies, demographic density (urban scale) and industrial infrastructure were important for determining the knowledge production in Brazil. Also, for the Brazilian microregions, Araújo and Garcia (2019) estimated a KPF using the SAR-Tobit method and found that higher levels of regional industrial R&D and academic research imply greater innovation measured by patents. Moreover, denser and diverse cities tend to present a better innovative performance what indicates Jacobian advantages for regional innovation in Brazil. Finally, Araújo and Garcia (2019) and Gonçalves, Oliveira and Almeida (2020) highlight the importance of interregional knowledge spillovers associated to innovative activities in Brazil.

Montenegro, Gonçalves and Almeida (2011) and Gonçalves, Matos and Araújo (2018) are the only papers in the literature that seek to estimate a KPF with a Dynamic Panel to search for path dependence in Brazil. However, the authors used the Generalized Method of Moment (GMM) estimator, different from the one proposed in this paper that also considered possible impacts from spillovers of knowledge production in *t-1* in neighbors' regions. Montenegro, Gonçalves and Almeida (2011) and Gonçalves, Matos and Araújo (2018) are the only papers in the literature (2011) did the research only for the microregions of São Paulo in the 1996-2003 period while Gonçalves, Matos and Araújo (2018) considered all microregions in Brazil in the 2000-2011 period. Both authors confirmed the importance of path dependence for the production of knowledge. In addition, Gonçalves, Matos and Araújo (2018) found evidences of knowledge spatial spillovers that reinforce the path-dependence process.

3 Empirical design

3.1 An expanded knowledge production function with path dependence and spatial spillovers

The present paper seeks to understand the knowledge production in Brazil, proposing an empirical design based on Fischer, Scherrngell and Reismann (2009) and Autant-Bernard and LeSage (2011). We initially propose the following equation that relates the knowledge produced by a given region according to the inputs needed to generate it

$$K_{it} = f(L_{it}, C_{it}, r_{it}) \tag{7}$$

where the regions are denoted by $i = 1, \ldots, N \in R \subset \mathbb{N}$, in which R is a natural well-order and upper bounded set of the regions; while t = 1, ..., N refers to the time period; L and C is the stock of labor and capital employed in knowledge production, respectively; and r_{it} are the other factors not explicitly considered that influence the creation of knowledge in the region i in period t. In addition, $f\colon K_{it} \rightarrow \mathbb{N}$ is a homogeneous function of degree one with decreasing returns of scale to capital (C) and labor (L): $\partial^2 K_{it}/\partial C_{it}^2 < 0$ and $\partial^2 K_{it}/\partial L_{it}^2 < 0$, respectively. Lastly, K_{it} is the amount of knowledge produced by the region i in period t. Assuming that $f\colon K_{it} \rightarrow \mathbb{N}$ is a Cobb-Douglas functional form, as in Fischer, Scherrngell and Reismann (2009),

$$K_{it} = f(L_{it}, C_{it}, r_{it}) = L_{it}^{\beta_1} C_{it}^{\beta_2} r_{it}^{\beta_3} \exp(\varepsilon_{it}) \tag{8}$$

where $\beta_{1,}$ β_2 are β_3 the elasticities of knowledge production in relation to the labor, capital and other fac3tors related to the creation of knowledge, respectively; all for region i in period t. Making a log transformation in (8) based on Jaffe (1989) and combining the empirical approach used by Fischer, Scherrngell and Reismann (2009) and Autant-Bernard and LeSage (2011),

$$\begin{aligned} \log(K_{it}) &= \rho \boldsymbol{W} \log(K_{it}) + \gamma\log(K_{it-1}) + \omega \boldsymbol{W} \log(K_{it-1}) + \beta_1 \log(C_{it-1}) \\ &+ \beta_2 \log(L_{it-1}) + \beta_3 \log(r_{it-1}) + \boldsymbol{WI}\tau + \epsilon_{ikt} \end{aligned} \tag{9}$$

Where $\boldsymbol{W}$ is a spatial lag matrix, $n \; x \; n$ which represents the spatial structural configuration between the regions; ρ and τ are parameters that capture the spatial interactions of variables log(k_{it-1}) and of the inputs and structural vector $\boldsymbol{I} = [\log(C_{it-1}) + \log(L_{it-1}) + \log(r_{it-1})]$. log(k_{it-1})represents the time lag of the dependent variable that seeks to capture the path dependence effects. $\boldsymbol{W}\log(K_{it-1})$ seeks to capture the spillover of knowledge production in t - 1 from region i on the creation of its neighbors j in t, according to a spatial configuration matrix represented by $\boldsymbol{W}$; γ is the coefficient that captures the force of path dependence on knowledge production in the region i while ω are coefficients that capture the space-time spillover between regions. To reduce potential endogeneity problems, we included the control variables in t - 1 since they impact knowledge production in t but are not influenced by innovation in t.

The Dynamic Spatial Panel Model approach, besides incorporating the spatial lag of the dependent variable, also incorporates a temporal dependent variable, in addition to a space-time lag of the dependent variable. Therefore, it is a methodology capable of empirically grasping the empirical design proposed in this article. The estimation of such a model follows the approach proposed by Yu, De Jong and Lee (2008), using a quasi-maximum likelihood estimator with fixed effects. The stationary region is satisfied with $|\rho| + |\gamma| + |\omega| < 1$ and we used the bias corrected estimator that consider the lower temporal dimension when compared to the spatial dimension of the data.

3.2 Data and variables

The proxy for knowledge production is the number of innovation patents created in the 558 microregions of Brazil for the 2005-2015 period. Patent deposit data were acquired in the Statistical Database of Intellectual Property (*Base de Dados Estatísticos sobre Propriedade Industrial - BADEPI*), generated by INPI (*Instituto Nacional de Propriedade Intelectual*). The patents are based on the first depositor's residence, which differs from a part of the literature that uses the inventor's residence. This choice reflects the discontinuation of the BADEPI (*Base de Dados Estatísticos sobre Propriedade Intelectual*) database that contained patents by inventor's residence only up to 2012.

We propose an empirical design based on microregion-level because the major part of the Brazilian municipalities did not present patents in the period, which could bias our results. The population data is from the *Instituto Brasileiro de Geografia e Estatística (IBGE)*. It is worth mentioning that all variables are from the 2005-2015 period. Then, we constructed an indicator of innovation patents per 100,000 inhabitants,

$$K_{it} = \frac{Innovation.P_{it}}{(Population_{it}/100.000)} \tag{10}$$

where K_{it} represents the indicator of patents per 100,000 inhabitants for microregion i in period t, which represents the *Knowledge Production* in the EKPF; I*nnovation.* P_{it} is the number of innovation patents; *Population*$_{it}$ is the size of the population of the microregion i in period t. Therefore, microregions with lower population gain more weight in knowledge production, when compared directly with those that have large populations. This enables to measure more effectively the knowledge productivity of these regions. The explanatory variables, inputs in the EKPF, are described in Table 1.

TABLE 1
Explanatory variables used in the Expanded Knowledge Production Function.

Variable	Description	Source
Knowledge Production	Innovation patents per 100,000 inhabitants.	INPI
Researchers	Number of researchers in public and private universities.	CAPES
Private R&D	Proportion of technical-scientific workers in total employment.	RAIS(1)
Higher Education	Portion of workers with higher education.	RAIS
Firms	Establishments with 2 to 500 employees.	RAIS
Large Firms	Proportion of companies with more than 500 employees	RAIS
Export	Exports of high and mid-high technology intensities products.	COMEX(2)
GDP	Per Capita Gross Domestic Product.	IBGE
Public Sci-Tech	Public expenditure on science and technology.	FINBRA(3)
Population Density	(Demographic Density) number of inhabitants per km².	IBGE

Source: research data.
(1) RAIS provides information about the Brazilian labor market.
(2) COMEX is a system for querying and extracting Brazilian foreign trade data.
(3) FINBRA is a system linked to the National Treasury that provides fiscal statistics for Brazilian municipalities.

Internal R&D performed in private companies, combined with those spent in universities and research institutes are key elements in the production of knowledge. In this context, we considered the number of researchers in public and private universities encompassed by Masters and PhD professionals working at universities available at CAPES (*Coordenação de Aperfeicoamento de Pessoal de Nível Superior*). It is worth mentioning that there is no data available on private R&D; therefore, the proportion of professionals employed in technical and scientific activities in the microregion was used as proxy, following Araújo, Cavalcante e Alves (2009).[3] According to Freitas *et al*. (2010) and Montenegro, Gonçalves and Almeida (2011), this variable is a suitable proxy due to the high correlation between them; therefore, it is the best variable to represent private R&D, given the lack of data in Brazil. The Higher Education is the proportion of workers that have a higher education in the workforce. Both the Higher Education and private R&D come from the RAIS (*Relação Anual de Informações Sociais*) from the Ministry of Economy. To capture

3 The professionals included are from the following areas: biotechnology, biomedical, engineers, researchers (in firms), mathematics and statistics professionals, computer systems analysts, physicists, chemists, space and atmosphere professionals and architects.

the role of public sector investments, we used the public expenditure in science and technology (Public Sci-Tech) available at the FINBRA – a report of information on expenses and revenues of each Brazilian municipality. This variable seeks to capture public efforts to incentive knowledge production in addition to the public expenditure in universities and research institutes.

On the other hand, the Population Density seeks to capture economies of agglomeration and Jacobian externalities. According to Jacobs (1969), geographical proximity makes possible an increase in the diffusion, transmission and exchange of ideas and information, both in their tacit and codified form, resulting in the amplification of knowledge production. This, for the most part, is linked to the urbanization of a certain locality, a result of the concentration of economic activities. In addition, several papers empirically corroborate the importance of agglomeration economics and Jacobian externalities for the innovative process. (GONÇALVES and ALMEIDA, 2009; FELDMAN and AUDRETSCH, 1999; GLAESER *et al.*, 1992). However, an excessive increase in population density may lead to agglomeration diseconomies, resulting in negative externalities, leading to a non-linear relationship with innovation. Such a fact can be captured with the inclusion of a term in the linear form and another in the quadratic form in KPF. If the linear is positive and the quadratic is negative, both significant, we have the presence of agglomeration diseconomies, acting together with the positive externalities. The overlapping of the two effects will depend on the magnitude of the population density, with negative externalities standing out in densely populated areas.

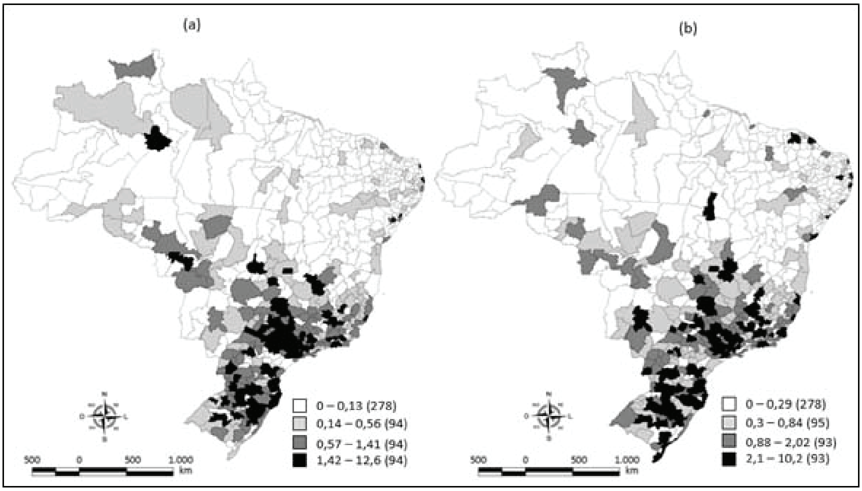
The Export variable captures the exports of high and mid-high technology intensities products and the data were made available on COMEX STAT (2021). In the high intensity products, we have the aerospace, pharmaceutical, computational, electronic and telecommunications sectors. On the other hand, in the mid-high intensity products, we have the electrical material; automotive vehicles; chemistry; railroad and transportation equipment; machinery and equipment. High and mid-high intensity sectors are both intensive in capital and technology, which can influence the innovation process and in the creation of knowledge, both directly and indirectly, through spillovers.

To capture the role of small (Firms) and large firms (Large Firms) in the knowledge production, we used the number of formal establishments with more than one employee per capita and the proportion of companies with more than 500 employees, respectively, from the RAIS (*Relação Anual de Informações Sociais*). The Firms variable seeks to capture the competitive structure of the local labor market since, according to Jacobs (1969) and Glaeser *et al.* (1992), a more competitive market produces greater incentive to innovation. On the other hand, the Large Firms variable tests if a regional structure with the presence of large companies could be more innovative, following Freitas, Gonçalves and Montenegro (2010). The GDP variable, representing the per capita Gross Domestic Product, seeks to control for the region's economic scale, which may influence the innovation process directly and/or indirectly by impacting its direct determinants. Finally, we also check for possible correlations between the variables (Appendix A) in order to avoid multicollinearity problems. From them, we can notice no extremely high correlations that could compromise the model estimation.

4. Spatial Distribution of Knowledge Production and its Determinants in Brazil

Innovation and technological improvements have important repercussions on the economic structure of a country, because they create new combinations of factors of production and generate alternative productive processes. A preliminary data analysis is performed in order to verify the dynamics of knowledge production in Brazil (see Figure 1).

FIGURE 1
Distribution of patents per 100,000 inhabitants in 2005 (a) and 2015 (b)



Source: INPI, data organized by the research

For 2005, the country presented 3740 innovation patents, while in 2015 this number rose to 5193, an increase of 38.85%. However, such growth does not occur homogeneously throughout the country. When considering regional levels, for example, a small number of microregions is responsible for most knowledge produced (patents) in the period. It is also evident the concentration of this innovative process, located mainly in the Center-South portion of the country, which was called by Gonçalves (2007) and Araújo and Garcia (2019), a North-South polarization regime. When comparing the initial period (a) versus the final (b), the production of patents per 100,000 inhabitants maintained the same spatial distribution and concentration. Therefore, the empirical evidences suggest the existence of an inertial process in the creation of patents, which indicates a potential path dependence for the production of knowledge in Brazil.

In addition, the spatial concentration in both periods is visible and to confirm we used the Exploratory Spatial Data Analysis (ESDA), which measures effects of spatial dependence and heterogeneity, association patterns (spatial clusters) and how the data are distributed. In particular, the Moran's I statistics calculate the spatial autocorrelation of patentes across regions and its coefficients are in Table 2; whose values were positive and statistically significant independent of the weight matrix applied. Thus, microregions with a high number of patents per 100,000 inhabitants tended to be surrounded by microregions with also high values.

TABLE 2
Moran's I for Patents (per 100,000 inhabitants) in 2005 and 2015

	Weights Matrix			
	Three Neighbors	**Five Neighbors**	**Seven Neighbors**	**Ten Neighbors**
Patents 2005	0,40(1)	0,38(1)	**0,41**(1)	0,39(1)
Patents 2015	0,34(1)	0,33(1)	**0,35**(1	0,33(1

Source: research data.
(1) Significance of 1%.
Note: Empirical Pseudo-significance based on 99999 random permutations.

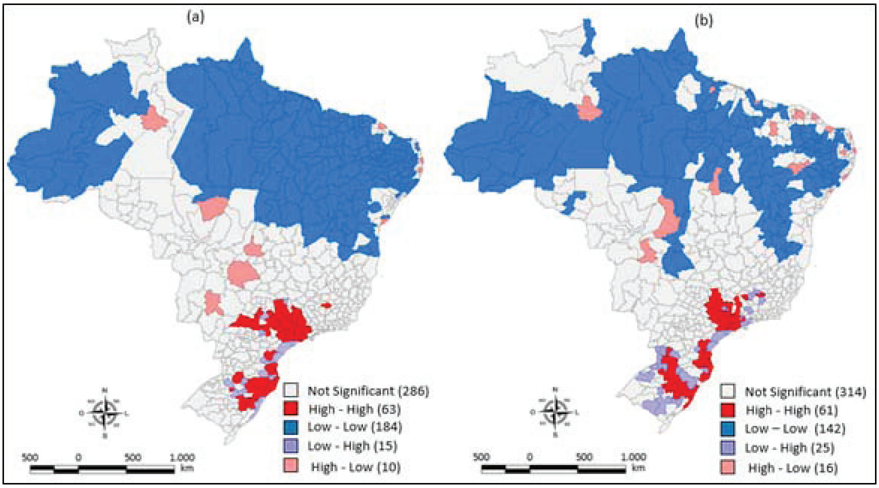
However, the Moran's I statistic only captures global autocorrelation, not identifying association at a local level. For this, we use the local Moran I (LISA), which is capable to capture local spatial autocorrelation and clusters that can be represented by four types: High-High (AA), Low-Low (BB), High-Low (AB) and Low-High (BA).[4] By using Lisa maps, we confirmed that there is a consolidation of

4 The most analyzed is the High-High cluster, which indicates that a region with a high value for the one variable is surrounded by regions with similar values.

two High-High (HH) clusters in the period (Figure 2), which represent microregions with high innovation surrounded by neighbors with equal innovative productivity.

The first HH cluster is located in the state of São Paulo and is configured as the largest spatial HH cluster due to both the number of microregions and the territorial extent of its spatial agglomeration. This polygon of high technological activity, in period one (a), extends from São Paulo and closer neighbors such as Guarulhos, Osasco and Santos, passing by Campinas, Piracicaba, São Carlos, Araraquara, Ribeirão Preto, to the north and northwest of the state, in Franca, Barretos and São José do Rio Preto. According to Montenegro, Gonçalves and Almeida (2011), the high technological development of the region reflects a diversified industrial base, as well as the presence of specialized sectors that encourage innovation.

FIGURE 2
LISA maps for patents per 100,000 inhabitants - 2005 (a) and 2015 (b)



Source: INPI, data organized by the research

However, in the second period (b), some microregions in the north and northwest were no longer significant, thus concentrating this cluster of high technological activity closer to microregions such as São Paulo and Campinas. According to Fernandes, Côrtes and Pinho (20004), the technology-based companies in São Paulo suffered from the country›s economic crisis, which possibly impacted their innovative capabilities. Finally, this HH cluster is in line with the evidences found by Gonçalves (2007), Montenegro and Betarelli Júnior (2008), Montenegro, Gonçalves and Almeida (2011), Araújo and Garcia (2019) and Góis Sobrinho and Azzoni (2016).

The second HH spatial cluster, in the first period (a), is located mainly in the eastern portion of the state of Santa Catarina (SC) and Rio Grande do Sul (RS). In SC, two local groups of the HH type stand out, one formed by Joinville, Blumenau and Itajaí and the other formed by Florianópolis, Tubarão and Criciúma. These regions have an important information technology sector, in addition to mechanical and electrical industries. The proximity of these clusters may indicate the presence of technological spillovers between the microregions. In the second period (b), the number of clusters of type HH was widened, consolidating the spatial cluster of the state, which starts at its border with RS, passing through the entire regional coast of SC up to Curitiba in the Paraná state, which is an important industrial complex of the south of Brazil, especially with regard to its automotive sector (GONÇALVES, 2007; GÓIS SOBRINHO; AZZONI, 2016).

The state of Rio Grande do Sul, in turn, presents most of the southern cluster in both periods considered. Moreover, from 2005 to 2015, there was an enlargement of this spatial agglomeration, extending from Porto Alegre and Caxias do Sul to Passo Fundo, also including adjacent regions such as Gramado, Montenegro, Guaporé and Não-Me-Toque, forming a corridor of technological development in the State. According to Araújo and Garcia (2019), these regions have physical and technical-scientific infrastructure suitable for knowledge production, with a skilled labor market and a dense industrial network.

Identified the basic spatial characteristics of the knowledge production in Brazil, the next step is to find its basic determinants. Table 3 brings the results of the Expanded Knowledge Production Function with path dependence and spatial spillovers, estimated with a Dynamic Spatial Panel method for the period from 2005 to 2015. Prior to the estimations, we applied the Hausman test to check the adequacy of the fixed effect estimation. The test rejected the null hypothesis that there is no systematic difference between the estimated coefficients[5]. In addition, we chose the spatial lag matrix that generated the lowest Akaike Information Criterion (AIC) for the benchmark model (Column IV), considering 3 to 100 k-neighbor weight matrixes, opting for the sixteen neighbors matrix. It is worth mentioning that we estimated the spatial models using the robust standard error to control for heteroscedasticity. Finally, to further support our empirical approach, we estimated the models by including gradually the control variables to check the robustness of the results and by estimating the model with a static fixed effect.

5 Chi^2 statistics: 75.41, with a p-value of 0,000.

TABLE 3
The Expanded Knowledge Production Function (EKPF) for Brazil

	(I)	(II)	(III)	(IV)	(V)	(VI)
Knowledge (y) t-1	0.1909[1]	0.1914[1]	0.1832[1]	0.1776[1]		
	(0.0261)	(0.0260)	(0.025)	(0.0256)		
W Knowledge (ρ) t	0.2230[1]	0.2146[1]	0.2169[1]	0.1804[1]	0.2115[1]	
	(0.0516)	(0.0513)	(0.0515)	(0.0512)	(0.0439)	
W Knowledge (ω) t-1	0.0666	0.0705	0.0773	0.0568		
	(0.0563)	(0.0567)	(0.0588)	(0.0590)		
Researchers	0.00028[1]	0.00027[1]	0.00026[1]	0.00026[1]	0.00032[1]	0.00031[1]
	(3.01E-05)	(2.94E-05)	(3.09E-05)	(3.16E-05)	(3.49E-05)	(3.62E-05)
Private R&D	2.98E-05	8.38E-05	3.65E-05	-1.5E-05	1.5E-05	-2.60E-05
	(8.11E-05)	(9.78E-05)	(9.13E-05)	(9.27E-05)	(7.79E-05)	(0.001)
Higher Education	0.7108	0.5095	0.4018	3.68E-01	0.3237	0.4460
	(0.4732)	(0.4074)	(0.3979)	(4.00E-01)	(0.3836)	(0.8001)
Firms		17.7533	22.0337	23.1248	22.8345	12.2605
		(17.7760)	(16.9684)	(17.3372)	(15.7688)	(7.6959)
Large Firms		1.9575	1.2205	7.66E-01	0.3305	0.3836
		(1.5968)	(1.4376)	(1.3200)	(1.3358)	(4.4725)
Population Density			0.0106[1]	0.0098[1]	0.0103[1]	0.0105[1]
			(0.0026)	(0.0002)	(0.0024)	(0.0015)
Population Density²			-9.8E-07[1]	-8.7E-07[1]	-9.1E-07[1]	-9.87E-07[1]
			(2.11E-07)	(2.40E-07)	(2.32E-07)	(2.41E-07)
Export				2.27E-05	2.8E-05	3.35E-05[2]
				(3.56E-05)	(3.6E-05)	(1.67E-05)
GDP				2.16E-06	1.73E-06	5.74E-06
				(3.32E-06)	(3.46E-06)	(4.46E-06)
Public Sci-Tech				-5.94E-10	-c5.79E-10	-3.40E-10
				(4.64E-10)	(4.29E-10)	(1.48E-10)
W Researchers	0.0003	0.0003	0.0002	0.0002	0.0003	
	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	
W Private R&D	-0.0008	0.0008	-0.0001	-0.0006	-0.0008	
	(0.0010)	(0.0015)	(0.0012)	(0.0012)	(0.0010)	
W Higher Education	6.7405[1]	2.9150	3.0411	1.1924	2.3395	
	(2.2159)	(2.2965)	(1.8571)	(1.8610)	(1.9129)	
W Firms		6.1025	0.3291	-71.8209[1]	-84.4852[1]	
		(21.6407)	(21.7288)	(27.7856)	(24.239)	

(continued)

TABLE 3
The Expanded Knowledge Production Function (EKPF) for Brazil

(continued)

	(I)	(II)	(III)	(IV)	(V)	(VI)
W Large Firms		-9.4615	-3.0222	-11.4896	-21.8387	
		(10.6336)	(9.4624)	(9.6881)	(9.5755)	
W Population Density			-0.0101	-0.0115	-0.0132	
			(0.0061)	(0.0061)	(0.0066)	
W Population Density²			1.29E-06	8.51E-07	1.10E-06	
			(1:21E-06)	(1.28E-06)	(1.28E-06)	
w GDP				7.64E-05(1)	7.54E-05(1)	
				(1.80E-05)	(1.78E-05)	
W Export				3.29E-05	-2.75E-06	
				(9.31E-05)	(9.37E-05)	
W Public Sci-Tech				6.42E-09	7.86E-09	
				(6.49E-09)	(6.53E-09)	
Observations	5022	5022	5022	5022	5580	5580
AIC	11572.40	11573.73	11524.99	11493.41	13011.55	13097.11
BIC	11637.62	11665.91	11629.99	11623.85	13130.84	13170.00
DSDM x DSAR: $\beta_6 = 0$						
X² Statistic	14,00	4.16	7.44	27.23	27.04	
p-value	0,0029	0.5266	0,3842	0.0013	0.0014	
DSDM x SEM: $\beta_6 = -\rho\beta_i$						
X² Statistic	15,84	6.23	6.77	27.53	26.61	
p-value	0,0012	0.2845	0,4536	0.0011	0.0016	

Source: research results.
(1) $\rho < 0.01$
(2) $\rho < 0.05$

To define which spatial model that best represents and captures the dynamics of the knowledge production in Brazil, we compared the Dynamic Spatial Autoregressive Model (DSAR) and the Dynamic Spatial Durbin Model (DSDM). The comparison is performed by testing the hypothesis that the spatially lagged independent variables (WX) are jointly non-significant. The null hypotheses (τ = 0) was rejected by the $\chi 2$ test, which presented a value of 27.23, significant at 1%, confirming that the Dynamic Spatial Durbin Model (DSDM) is the best model. To check the importance of the path dependence in knowledge production, capture by the temporal coefficient, we compared the results from the dynamic model with a static Spatial Durbin Model (SDM). By using a Likelihood-Ratio (LR) test, which resulted in a value of 50.86, significant at 1%, we confirmed the importance of path dependence to understand the knowledge production in Brazil. Finally, we test the Dynamic Spatial Durbin Model (DSDM) against the Spatial Error Model (SEM) to further support our empirical approach. In other words, if $\beta 6 = -\rho\beta i$ holds, then, the Spatial Error Model is the best model. The $\chi 2$ statistic presented a value of 27.53, significant at 1%, reinforcing the Dynamic Spatial Durbin Model (DSDM) as the best model to understand knowledge production in Brazil.

Among the main results, we highlight that the time-lagged dependent variable, K_{t-1}, presented a coefficient (γ) statistically significant at 1%, with a positive impact on knowledge production. Therefore, the assumption of Fischer, Scherrngell and Reismann (2009) that the production of knowledge in is positively influenced by the quantity produced in $t-1$ was true for the microregions of Brazil. In others words, the innovation in the country has an inertial component, assuming a pre-established trajectory, or path dependence, as in Arthur (1989). This fact explains the dynamics presented by the innovative activity in Figures 1 and 2, where it is evident that the regions that were the largest producers of knowledge in 2005 continued to be in 2015, with few visible changes.

On the other hand, the coefficient that aims to capture the knowledge spillovers in $t, (\rho)$, presented statistical significance while the coefficient that captures spillovers from $t-1$, (ω), did not. Therefore, only the present spatial spillovers component, WK_{it}, is relevant to explain knowledge production in Brazil, indicating that regions with high innovation influence positively their neighbors, a fact that can partially explain the spatial concentration of knowledge in Figures 1 and 2. The spatial concentration phenomena occur because certain activities are agglomerated in a given locality due to the presence of attractive (centripetal) forces. This result corroborates Gonçalves and Almeida (2009), Gonçalves, Matos and Araújo (2018), Araújo and Garcia (2019) who found significant knowledge spillovers in the country.

Therefore, considering the nature of this model, we cannot interpret the coefficients from Table 3 directly due to the presence of spatial spillovers and temporal persistence that causes indirect and long run effects, which are presented in Table 4. Among the results, we highlight the number of university researchers in the microregions, which obtained a positive and statistically significant direct and long run coefficients. This variable is part of the labor force, L_t, employed in the search for new ideas, being intrinsically related to scientific and technological research, which, according to Nelson (1996), are important vectors of knowledge production and innovation.

However, other components of the labor input in the EKPF, such as Private R&D (which had as proxy, the professionals related to this activity) and workers with higher education are not significant, which indicate a low return, in terms of innovation, for R&D in the Brazilian firms. In particular, the statistical insignificance of Private R&D contrasts with the empirical evidences of Araújo and Garcia (2019) and Gonçalves, Oliveira and Almeida (2020), but is in line with Montenegro *et al.* (2011). One potential explanation for this empirical result is that the Araújo and Garcia (2019) and Gonçalves, Oliveira and Almeida (2020) did not control for path dependence that prevails in knowledge production in Brazil. On the other hand, Montenegro, Gonçalves and Almeida (2011), by controlling for an inertial component in the innovation process, found no statistically significant evidence that private R&D impacts independently the knowledge production.

TABLE 4

Direct, indirect and total effects on the short and long run

	SR_Direct	SR_Indirect	SR_Total	LR_Direct	LR_Indirect	LR_Total
Researchers	0.0002(1)	0.0002	0.0005(2)	0.0003(1)	0.0004	0.0007(2)
	(3.29E-05)	(0.0002)	(0.0002)	(0.0001)	(0.0003)	(0.0003)
Private R&D	-1.8E-05	-0.0007	-0.0007	-2.9E-05	-0.0010	-0.0010
	(0.0010)	(0.0071)	(0.0073)	(0.0013)	(0.0105)	(0.0108)
Higher Education	0.3589	0.9385	1.2974	0.4552	1.4656	1.9209
	(0.7304)	(3.3231)	(3.4044)	(0.9080)	(4.8979)	(5.0484)
Firms	20.5378	-81.6009(1)	-61.0631(1)	24.6862	-115.087(1)	-90.4009(1)
	(14.8738)	(24.0151)	(19.9453)	(18.3407)	(33.5954)	(29.7090)
Large Firms	-0.0223	-28.6704	-28.6927	-0.3176	-42.1757	-42.4934
	(4.4346)	(16.2231)	(16.9003)	(5.5166)	(24.0005)	(25.1424)
Population Density	0.0083(1)	-0.0213(1)	-0.0130(1)	0.0101(1)	-0.0293(1)	-0.01926
	(0.0015)	(0.0064)	(0.0066)	(0.0019)	(0.0095)	(0.0098)
Population Density^2	-6.6E-07(1)	2.4E-06	1.8E-06	-8.0E-07(1)	3.4E-06	2.6E-06
	(2.3E-07)	(1.3E-06)	(1.3E-06)	(2.9E-07)	(1.9E-06)	(2.0E-06)
Export	2.0E-05	-2.8E-05	-8.3E-06	2.4E-05	-3.6E-05	-1.2E-05
	(1.6E-05)	(7.4E-05)	(7.8E-05)	(2.0E-05)	(0.0001)	(0.0001)
GDP	3.0E-06	8.9E-05(1)	9.2E-05(1)	4.6E-06	1.3E-04(1)	1.4E-04(1)
	(4.3E-06)	(1.9E-05)	(1.9E-05)	(5.3E-06)	(2.8E-05)	(2.9E-05)
Public Sci-Tech	-1.0E-09	1.3E-09(2)	1.2E-08	-1.1E-09	1.8E-09	1.7E-08
	(1.3E-09)	(6.4E-09)	(6.6E-09)	(6.6E-09)	(9.5E-09)	(9.9E-09)

Source: research results.
(1) $\rho < 0.01$
(2) $\rho < 0.05$

Another important determinant of knowledge production in Brazil is the population density in both linear and quadratic form. They were significant, for the direct effects, at the 1% level and with expected signs, that is, with positive and negative coefficients, respectively. This relationship confirms for Brazil the hypothesis of Jacobs (1969), Krugman (1991) and Griliches (1992) on the importance of agglomeration economies and urbanization in the production of knowledge. Therefore, the increase of population density in a certain locality of the country is able to facilitate the diffusion, transmission and exchange of ideas among the economic agents, factors capable of increasing the returns of the knowledge produced due to the positive externalities. However, as indicated by the negative sign of the squared coefficient, these benefits of agglomeration occur until a certain urban scale, from which they begin to act as inhibitors in the production of knowledge. Gonçalves and Almeida (2009) and Araújo and Garcia (2019) also found a similar result for Brazil. Therefore, this work corroborates the importance of Jacobian externalities and agglomeration economies to explain the innovative process in the Brazilian microregions. However, it is worth mentioning that population density also has significant negative spillovers on neighbors which, in combination to the direct effects, makes its long-term impact null.

The microregions economic scale, captured by the per capita GDP, presented significant positive indirect effect on neighbors, enhancing its knowledge production. In other words, increased economic scale generates externalities that create incentives to the productive innovation process. Many factors can explain this phenomenon. For example, a higher GDP translates to a higher demand from the consumer side, which, in turn, allows gains on the production side due to decreasing costs in the presence of economies of scale and due to reducing risks from investments on innovation.

The large firms do not impact the knowledge production, which may reflect the significant presence of multinationals among large firms, where the innovation process takes place especially at the headquarters, along with market concentration and low competition among oligopolies. On the other hand, the number of firms generates a significant negative indirect effect. Finally, the Export of high and mid-high technology intensive products and public expenditures on science and technology are not statistically significant, which highlight concerns due to their usual importance in the literature.

5. Final remarks

This paper sought to propose an empirical approach for an Expanded Knowledge Production Function (EKPF) based on theoretical propositions that encompass both path dependence and spatial spillovers. In addition, we search for evidence in Brazil with a Dynamic Spatial Panel Data approach, using microregions as a basic geographic cut. The basic purpose was to identify the determinants of knowledge production in the country as well as its temporal evolution, using innovation patents as proxy in the period from 2005 to 2015.

An important evidence found regarding the distribution of knowledge production is the existence of a North-South disparity for the innovative activity in Brazil, with São Paulo and the South of Brazil having the two largest high-high clusters in the country. We also confirmed the importance of path dependence phenomenon and of spatial spillovers to understand knowledge production. In other words, the distribution of the innovative activity in the country presents an inertial element, with regions that started the creation of knowledge earlier having an advantage in their technological development. Hence, the unchanged North-South disparity for innovation in the country may be due to the presence of a path dependence process, which makes it difficult to reduce regional differences. In addition, spatial spillovers from knowledge reinforce the spatial agglomeration, since it influences the production of knowledge in neighbors' regions. Such evidence indicates the importance of targeted public policies for some localities to be able to initiate a technological development trajectory, without which they will not reach the leading regions in knowledge production.

In this context, we considered the indirect effects from spatial spillovers and the long run effects from path dependence, which improved the results interpretation. Among the main empirical evidences, we highlight the positive long run impact from the number of university researchers, a proxy for R&D conducted by universities, an important part of the labor input, , in the knowledge production function. On other hand, the number of researchers in the universities are important vectors of knowledge production, which reinforces its role in the innovation process. In addition, the population density presented an inverted "U" influence with innovation in the short run, capturing a nonlinear relation from agglomeration externalities. Put another way, these benefits of agglomeration occur until a certain urban scale, from which they begin to act as inhibitors in the production of knowledge. This fact corroborates the importance of urbanization and of the urban scale for Brazilian technological activity. However, population density presented negative spillovers on neighbors, which makes its long-term impact null.

Finally, we found empirical evidences that increased economic scale generates externalities that create incentives to the productive innovation process, possibly reflecting a higher demand from the consumer side and decreasing costs and investment risk on the supply side. From these evidences, we identified the key elements that boost innovation activity and the production of knowledge in Brazil. However, the lack of significance of the other factors included in the EKPF leaves open the reasons why these variables are not relevant for the country's knowledge production. As an example, we can mention the cases of private R&D, local public investments in science and technology and exports of high and mid-high technology intensive products that are expected to assist in the production of knowledge, but have not behaved as expected, evidencing a malfunction of these elements, a fact that deserves attention by public agents and researchers.

References

ALBUQUERQUE, E. da M. e; SIMÕES, R.; BAESSA, A.; CAMPOLINA, B.; SILVA, L. A. A distribuição espacial da produção científica e tecnológica brasileira: uma descrição de estatística de produção local de patentes e artigos científicos. *Revista Brasileira de Inovação*, v. 1, n. 2, p. 225-251, 2002.

ARAÚJO, V.D.C., GARCIA, R. Determinants and spatial dependence of innovation in Brazilian regions: evidence from a Spatial Tobit Model. *Nova Economia*, v. 29, n. 2, p. 375-400, 2019.

ARAÚJO, B.C.; CAVALCANTE, L.R.; ALVES, P. 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, v. 5, p. 16-21, 2009.

ARTHUR, B. Competing technologies, increasing returns and lock in by historical events. *The Economic Journal*, v. 99, n. 394, p. 116-131, 1989.

AUTANT-BERNARD, C. The Geography of Knowledge Spillovers and Technological Proximity. *Economics of Innovation and New Technology*, v. 10, n. 4, p. 237-254, 2001.

AUTANT-BERNARD, C.; LESAGE, J.P. Quantifying Knowledge Spillovers Using Spatial Econometric Models. *Journal of Regional Science*, v. 51, n. 3, p. 471-496, 2011.

AUTANT-BERNARD, C.; MAIRESSE, J.; MASSARD, N. Spatial Knowledge Diffusion through Collaborative Networks. *Papers in Regional Science*, v. 86, n. 3, p. 341-350, 2007.

BARROS, P.H.B.; FREITAS JUNIOR, A.M.; RAIHER, A.P.; STEGE, A.L. Distribuição das patentes municipais do Sul do Brasil: uma análise espacial, com ênfase nos efeitos locais. *Revista Brasileira de Inovação*, v. 18, n. 1, p. 9-36, 25 abr. 2019.

COMEX STAT. Estatísticas do Comércio Exterior Brasileiro, 2021. Available in: http://comexstat.mdic.gov.br/. Accessed on: 15 Nov. /2021.

EJERMO, O. Technological diversity and Jacobs' externality hypothesis revisited. *Growth and Change*, v. 36, n. 2, p. 167-195, 2005.

FELDMAN M.P.; AUDRETSCH D.B. Innovation in cities: science-based diversity, specialization and localized competition. *European Economic Review*, v. 43, n. 2, p. 409-429, 1999.

FERNANDES, A.C.; CÔRTES, M.R.; PINHO, M. Caracterização das pequenas e médias empresas de base tecnológica em São Paulo: uma análise preliminar. *Economia e Sociedade*, v. 13, n. 1 (22), p. 151-173, 2004.

FISCHER, M.; SCHERRNGELL, T.; REISMANN, M. Knowledge Spillovers and Total Factor Productivity: Evidence Using a Spatial Panel Data Model. *Geographical Analysis*, v. 41, n. 2, p. 204-220, 2009.

FREEMAN, C. Technical innovation, diffusion, and long cycles of economic development. *In:* VASKO, T. (ed.). *The long-wave debate*. Springer, Berlin: Heidelberg, 1988. p. 295-309.

FREITAS JÚNIOR, A.M.; BARROS, P.H.B; STEGE, A.L; SANTOS, C.V; HILGEMBERG, C.M.A.T. Distribuição espacial da inovação na região Sul do Brasil de 2005 a 2015, a partir da análise exploratória de dados espaciais. *Revista de Desenvolvimento Econômico – RDE,* ano XXIII, v. 1, n. 48, p. 31-59, abr. 2021.

FREITAS, M.V.; GONÇALVES, E.; MONTENEGRO, R.L.G. Desigualdade tecnológica, convergência espacial e transbordamentos: uma análise por estados brasileiros (1990-2001). *Revista Brasileira de Estudos Regionais e Urbanos*, v. 4, n.2, p. 1-21, 2010.

GLAESER E.L.; KALLAL, H.D.; SCHEINKMAN, J.A.; SHLEIFER, A. Growth in cities. *Journal of Political Economy*, v. 100, n. 6, p. 1126-1152, 1992.

GÓIS SOBRINHO, E.M.G.; AZZONI, C.R. Potencial inovativo da indústria nas regiões brasileiras. *Revista Brasileira de Inovação*, v. 15, n. 2, p. 275-304, jul./dez. 2016.

GONÇALVES, E. A distribuição espacial da atividade inovadora brasileira: uma análise exploratória. *Estudos Econômicos*, v. 37, n. 2, p. 405-433, 2007.

GONÇALVES, E.; ALMEIDA, E. S. Innovation and Spatial Knowledge Spillovers: Evidence from Brazilian Patent Data. *Regional Studies*, v. 43, n. 4, p. 513-528, May 2009.

GONÇALVES, E.; MATOS, C.M.; ARAÚJO, I.F. Path-Dependent Dynamics and Technological Spillovers in the Brazilian Regions. *Applied Spatial Analysis and Policy*, p. 1-25, 2018.

GONÇALVES, E.; OLIVEIRA, P.M.; ALMEIDA, E. Spatial determinants of inventive capacity in Brazil: the role of inventor networks. *Spatial Economic Analysis*, v. 15, n. 2, p. 186-207, 2020.

GRILICHES, Z. Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics*, v. 10, n. 1, p. 92-116, 1979.

GRILICHES, Z. *R&D, Patents, and Productivity.* Cambridge, MA: National Bureau of Economic Research, 1984.

GRILICHES, Z. Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, v. 28, n. 4, p. 1661-1707, 1992.

JACOBS, J. *The Economies of Cities*. New York: Random House, 1969.

JAFFE, A.B. Real effects of academic research. *American Economic Review*, v. 79, n. 5, p. 957-970, 1989.

KRUGMAN, P. *Geography and trade*. Cambridge, MA: MIT Press, 1991.

LUCAS, R.E. On the mechanics of economic development. *Journal of Monetary Economics*, v. 22, n. 1, p. 3-42, 1988.

MIRANDA, P.; ZUCOLUTO, G. Conhecimento com perfil inovador nas infraestruturas científicas e tecnológicas no Brasil. *Radar*, n. 37, IPEA, 2015.

MONTENEGRO, R.L.G.; BETARELLI JUNIOR, A. Análise e investigação dos fatores determinantes da inovação nos municípios de São Paulo. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 2, n. 2, p. 7-31, 2008.

MONTENEGRO, R.L.; GONÇALVES, E.; ALMEIDA, E. Dinâmica Espacial e Temporal da Inovação no Estado de São Paulo: Uma Análise das Externalidades de Diversificação e Especialização. *Estudos Econômicos*, v. 41, n. 4, out./dez. 2011.

NELSON, R. *Sources of Economic Growth*. Cambridge, Mass.: Harvard University, 1996.

NORTH, D. *Structure and Change in Economic History*. New York: W. W. Norton, 1981.

OLIVEIRA, P.M.; GONÇALVES, E.; ALMEIDA, E. Existe convergência de patenteamento no Brasil? *Revista Brasileira de Inovação*, Campinas (SP), v. 15, n. 2, p. 335-364, jul./dez. 2016.

PARENT, O.; LESAGE, J. Using the Variance Structure of the Conditional Autoregressive Specification to Model Knowledge Spillovers. *Journal of Applied Econometrics*, v. 23, n. 2, p. 235-256, 2008.

RODRIGUEZ, R.S., GONÇALVES, E. Hierarquia e concentração na distribuição regional brasileira de invenções por tipos de tecnologias. *Revista Brasileira de Inovação*, v. 16, n. 2, p. 225-266, 2017.

ROMER, P. Endogenous technological change. *Journal of Political Economy*, v. 98, n. 5, 1990.

ROMER, P. *Two Strategies for Economic Development:* Using Ideas and Producing Ideas. *In*: PROCEEDINGS OF THE WORLD BANK ANNUAL CONFERENCE ON DEVELOPMENT ECONOMICS. Washington, D.C, 1993.

SOLOW, R.M. A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, v. 70, n. 1, p. 65-94, 1956.

YU, J.; DE JONG, R.; LEE, L.F. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both N and T are large. *Journal of Econometrics*, v. 146, n. 1, p. 118-134, 2008.

Author's contribution

A. Literature review and problematization: Pedro Henrique Batista de Barros

B. Data collection and statistical analysis: Pedro Henrique Batista de Barros and Adirson Maciel de Freitas Júnior

C. Preparation of figures and tables: Pedro Henrique Batista de Barros and Adirson Maciel de Freitas Júnior

D. Manuscript development: Pedro Henrique Batista de Barros

E. Bibliography selection: Pedro Henrique Batista de Barros

Conflict of interest: the authors declare that there is no conflict of interest

Funding: the authors declare that there is no funding

Appendix

Correlation for the variables use in the econometric estimation

	R&D	Research	Firm	Sci. Tech	High. Ed	L. Firm	Ex-ports	GDP	Pop. Dens
R&D	1.000								
Research	-0.002	1.000							
Firms	-0.046	0.054	1.000						
Sci. Tech.	0.000	0.090	0.025	1.000					
Higher Ed.	0.076	0.159	0.215	0.070	1.000				
Large Firms	0.075	-0.030	-0.544	-0.007	-0.056	1.000			
Exports	-0.004	0.012	0.056	0.013	0.049	-0.027	1.000		
GDP	-0.027	0.087	0.529	0.059	0.237	-0.244	0.306	1.000	
Pop.Dens	0.001	0.358	0.053	0.239	0.300	-0.023	0.033	0.204	1.000

Source: research results.